

# PROCESS, SYSTEM AND COMPUTER READABLE MEDIUM FOR PULMONARY NODULE DETECTION USING MULTIPLE-TEMPLATES MATCHING

**Patent number:** JP2004517660 (T)

**Publication date:** 2004-06-17

**Inventor(s):**

**Applicant(s):**

**Classification:**

**- international:** **A61B5/00; A61B5/055; A61B6/00; A61B6/03; A61B8/00; G01R33/32; G06T1/00; G06T7/00; A61B5/00; A61B5/055; A61B6/00; A61B6/03; A61B8/00; G01R33/32; G06T1/00; G06T7/00;** (IPC1-7): A61B5/00; A61B5/055; A61B6/00; A61B6/03; A61B8/00; G01R33/32; G06T1/00; G06T7/00

**- european:** G06T7/00B2

**Application number:** JP20020545549T 20011121

**Priority number(s):** US20000716335 20001121; WO2001US43146 20011121

**Also published as:**

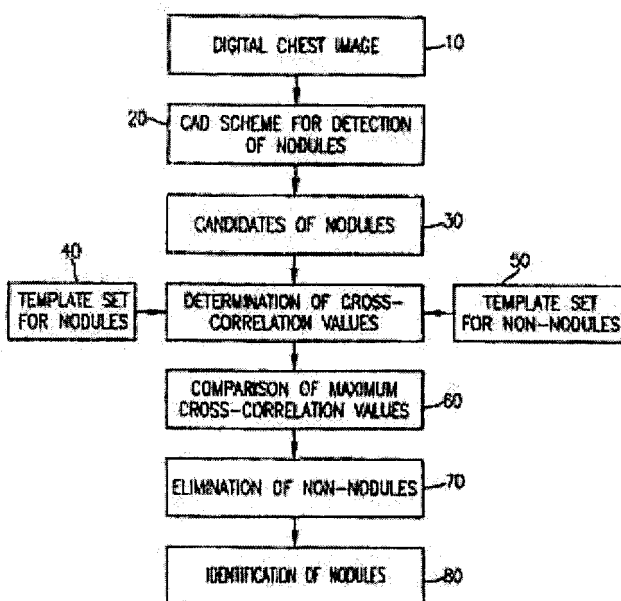
WO0243562 (A2)  
WO0243562 (A3)  
US6470092 (B1)  
US2003002723 (A1)  
US6683973 (B2)

more >>

Abstract not available for JP 2004517660 (T)

Abstract of correspondent: **WO 0243562 (A2)**

A method to determine whether a candidate abnormality in a medical digital image is an actual abnormality, a system which implements the method, and a computer readable medium which stores program steps to implement the method, wherein the method includes obtaining a medical digital image including a candidate abnormality; obtaining plural first templates and plural second templates respectively corresponding to predetermined abnormalities and predetermined non-abnormalities; comparing the candidate abnormality with the obtained first and second templates to derive cross-correlation values between the candidate abnormality and each of the obtained first and second templates; determining the largest cross-correlation value derived in the comparing step and whether the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates or with the second templates; and determining the candidate abnormality to be an actual abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates and determining the candidate abnormality to be a non-abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the second templates. An actual abnormality is similarly classified as malignant or benign based on further cross-correlation values obtained by comparisons with additional templates corresponding to malignant and benign abnormalities.



Data supplied from the **espacenet** database — Worldwide

FC

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-517660

(P2004-517660A)

(43) 公表日 平成16年6月17日(2004.6.17)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
A 61 B 5/00	A 61 B 5/00 G	4 C 0 9 3
A 61 B 5/055	A 61 B 6/03 3 6 0 D	4 C 0 9 6
A 61 B 6/00	A 61 B 8/00	4 C 6 0 1
A 61 B 6/03	G 0 6 T 1/00 2 9 0	5 B 0 5 7
A 61 B 8/00	G 0 6 T 7/00 3 0 0 D	5 L 0 9 6
審査請求 未請求 予備審査請求 有 (全 87 頁) 最終頁に続く		

(21) 出願番号 特願2002-545549 (P2002-545549)  
 (86) (22) 出願日 平成13年11月21日 (2001.11.21)  
 (85) 翻訳文提出日 平成15年5月20日 (2003.5.20)  
 (86) 国際出願番号 PCT/US2001/043146  
 (87) 国際公開番号 W02002/043562  
 (87) 国際公開日 平成14年6月6日 (2002.6.6)  
 (31) 優先権主張番号 09/716,335  
 (32) 優先日 平成12年11月21日 (2000.11.21)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 502261163  
 ザ・ユニバーシティ・オブ・シカゴ  
 アメリカ合衆国、イリノイ州 60637  
 シカゴ、イー・エリス・アベニュー 5  
 801  
 (74) 代理人 100058479  
 弁理士 鈴江 武彦  
 (74) 代理人 100091351  
 弁理士 河野 哲  
 (74) 代理人 100088683  
 弁理士 中村 誠  
 (74) 代理人 100108855  
 弁理士 蔵田 昌俊  
 (74) 代理人 100075672  
 弁理士 峰 隆司

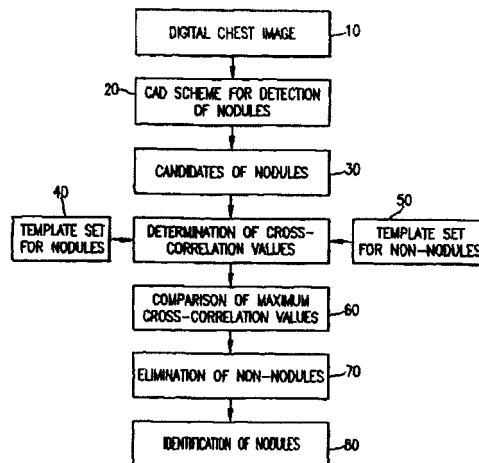
最終頁に続く

(54) 【発明の名称】 多重テンプレートマッチングを使用する肺小結節検出のためのプロセス、システムおよびコンピュータ読み取り可能な媒体

## (57) 【要約】

【課題】 多重テンプレートマッチングを使用する肺小結節検出のためのプロセス、システムおよびコンピュータ読み取り可能な媒体。

【解決手段】 医用デジタル画像の候補異常が真の異常であるかどうかを決定する方法、該方法を実現するシステム、および該方法を実現するプログラムステップを格納するコンピュータ読み取り可能な媒体。該方法は、候補異常を含む医用デジタル画像を得ること、予め定められた異常および予め定められた非異常にそれぞれ対応する複数の第1テンプレートおよび複数の第2テンプレートを得ること、候補異常を得られた第1および第2テンプレートと比較して、候補異常と得られた第1および第2テンプレートの各々との間の相互相関値を導出すること、比較工程で導出された最大相互相関値を決定し、かつ最大相互相関値が候補異常を第1テンプレートまたは第2テンプレートのどちらと比較することによって得られたかを決定すること、最大相互相関値が候補異常を第1テンプレートと比較することによって生成された場合には候補異常を真の異常と決定し、最大相互



## 【特許請求の範囲】

## 【請求項 1】

医用デジタル画像の候補異常が真の異常であるかどうかを決定する方法において、  
予め定められた異常および予め定められた非異常にそれぞれ対応する複数の第 1 テンプレートおよび複数の第 2 テンプレートを得ることと、  
前記医用デジタル画像の候補異常を前記得られた第 1 および第 2 テンプレートと比較して、前記候補異常と前記得られた第 1 および第 2 テンプレートの各々との間の相互相関値を導出することと、  
前記比較工程で導出された最大相互相関値を決定し、かつ前記最大相互相関値が前記候補異常を前記第 1 テンプレートまたは前記第 2 テンプレートのどちらと比較することによって生成されたかを決定することと、  
前記最大相互相関値が前記候補異常を前記第 1 テンプレートと比較することによって生成された場合には、前記候補異常を真の異常と決定し、前記最大相互相関値が前記候補異常を前記第 2 テンプレートと比較することによって生成された場合には、前記候補異常を非異常と決定することと、  
を具備する改良された方法。

10

## 【請求項 2】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記候補異常よりサイズが小さい複数の第 1 テンプレートおよび複数の第 2 テンプレートを得ることを具備する請求項 1 に記載の方法。

20

## 【請求項 3】

前記比較工程が、  
前記候補異常を前記得られた第 1 および第 2 テンプレートと比較して、前記候補異常と前記得られた第 1 および第 2 テンプレートの各々との間の相互相関値を導出することと、  
前記第 1 および第 2 テンプレートを前記候補異常に対してシフトさせて、前記候補異常と前記シフトされた第 1 および第 2 テンプレートの各々との間の相互相関値を導出することとを具備する請求項 2 に記載の方法。

## 【請求項 4】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの鏡像である追加第 1 および第 2 テンプレートを生成することを具備する請求項 1 に記載の方法。

30

## 【請求項 5】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートのスケーリングされたバージョンである追加第 1 テンプレートを生成することを具備する請求項 1 に記載の方法。

## 【請求項 6】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの回転バージョンである追加第 1 および第 2 テンプレートを生成することを具備する請求項 1 に記載の方法。

## 【請求項 7】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの鏡像である追加第 1 および第 2 テンプレートを生成することを具備する請求項 2 に記載の方法。

40

## 【請求項 8】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートのスケーリングされたバージョンである追加第 1 テンプレートを生成することを具備する請求項 2 に記載の方法。

## 【請求項 9】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの回転バージョンで

50

ある追加第 1 および第 2 テンプレートを生成することを具備する請求項 2 に記載の方法。

【請求項 1 0】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの鏡像である追加第 1 および第 2 テンプレートを生成することを具備する請求項 3 に記載の方法。

【請求項 1 1】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートのスケーリングされたバージョンである追加第 1 テンプレートを生成することを具備する請求項 3 に記載の方法。

【請求項 1 2】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの回転バージョンである追加第 1 および第 2 テンプレートを生成することを具備する請求項 3 に記載の方法。

【請求項 1 3】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートのスケーリングされたバージョンである追加第 1 テンプレートを生成することを具備する請求項 7 に記載の方法。

【請求項 1 4】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの回転バージョンである追加第 1 および第 2 テンプレートを生成することを具備する請求項 7 に記載の方法。

【請求項 1 5】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートのスケーリングされたバージョンである追加第 1 テンプレートを生成することを具備する請求項 1 0 に記載の方法。

【請求項 1 6】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
前記得られた複数の第 1 テンプレートおよび複数の第 2 テンプレートの回転バージョンである追加第 1 および第 2 テンプレートを生成することを具備する請求項 1 0 に記載の方法。

【請求項 1 7】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
候補第 1 テンプレートおよび候補第 2 テンプレートを得ることと、  
前記候補第 1 テンプレートを少なくとも 1 つの既知の非異常構造と比較し、前記少なくとも 1 つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第 1 テンプレートだけを第 1 テンプレートとして保持することと、  
前記候補第 2 テンプレートを少なくとも 1 つの既知の異常構造と比較し、前記少なくとも 1 つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第 2 テンプレートだけを第 2 テンプレートとして保持することとを具備する請求項 1 に記載の方法。

【請求項 1 8】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、  
候補第 1 テンプレートおよび候補第 2 テンプレートを得ることと、  
前記候補第 1 テンプレートを少なくとも 1 つの既知の非異常構造と比較し、前記少なくとも 1 つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第 1 テンプレートだけを第 1 テンプレートとして保持することと、  
前記候補第 2 テンプレートを少なくとも 1 つの既知の異常構造と比較し、前記少なくとも 1 つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第 2 テンプレートだけを第 2 テンプレートとして保持することとを具備する請求項 2 に記載の方法。

【請求項 1 9】

前記複数の第 1 テンプレートおよび複数の第 2 テンプレートを得る工程が、

候補第1テンプレートおよび候補第2テンプレートを得ることと、  
前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、  
前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項3に記載の方法。

【請求項20】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、  
候補第1テンプレートおよび候補第2テンプレートを得ることと、  
前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、  
前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項16に記載の方法。

【請求項21】

予め定められた悪性異常および予め定められた良性異常にそれぞれ対応する複数の第3テンプレートおよび複数の第4テンプレートを得ることと、  
真の異常を前記得られた第3および第4テンプレートと比較して、前記真の異常と前記得られた第3および第4テンプレートの各々との間の相互相関値を導出することと、  
前記比較ステップで導出された最大相互相関値を決定し、かつ前記最大相互相関値が前記真の異常を前記第3テンプレートまたは前記第4テンプレートのどちらと比較することによって生成されたかを決定することと、  
前記最大相互相関値が前記真の異常を前記第3テンプレートと比較することによって生成された場合には、前記真の異常を悪性異常と分類し、前記最大相互相関値が前記真の異常を前記第4テンプレートと比較することによって生成された場合には、前記真の異常を良性異常と分類することとをさらに具備する請求項1に記載の方法。

【請求項22】

予め定められた悪性異常および予め定められた良性異常にそれぞれ対応する複数の第3テンプレートおよび複数の第4テンプレートを得ることと、  
真の異常を前記得られた第3および第4テンプレートと比較して、前記真の異常と前記得られた第3および第4テンプレートの各々との間の相互相関値を導出することと、  
前記比較ステップで導出された最大相互相関値を決定し、かつ前記最大相互相関値が前記真の異常を前記第3テンプレートまたは前記第4テンプレートのどちらと比較することによって生成されたかを決定することと、  
前記最大相互相関値が前記真の異常を前記第3テンプレートと比較することによって生成された場合には、前記真の異常を悪性異常と分類し、前記最大相互相関値が前記真の異常を前記第4テンプレートと比較することによって生成された場合には、前記真の異常を良性異常と分類することとをさらに具備する請求項2に記載の方法。

【請求項23】

予め定められた悪性異常および予め定められた良性異常にそれぞれ対応する複数の第3テンプレートおよび複数の第4テンプレートを得ることと、  
真の異常を前記得られた第3および第4テンプレートと比較して、前記真の異常と前記得られた第3および第4テンプレートの各々との間の相互相関値を導出することと、  
前記比較ステップで導出された最大相互相関値を決定し、かつ前記最大相互相関値が前記真の異常を前記第3テンプレートまたは前記第4テンプレートのどちらと比較することによって生成されたかを決定することと、  
前記最大相互相関値が前記真の異常を前記第3テンプレートと比較することによって生成された場合には、前記真の異常を悪性異常と分類し、前記最大相互相関値が前記真の異常

10

20

30

40

50

を前記第4テンプレートと比較することによって生成された場合には、前記真の異常を良性異常と分類することとをさらに具備する請求項3に記載の方法。

【請求項24】

予め定められた悪性異常および予め定められた良性異常にそれぞれ対応する複数の第3テンプレートおよび複数の第4テンプレートを得ることと、  
真の異常を前記得られた第3および第4テンプレートと比較して、前記真の異常と前記得られた第3および第4テンプレートの各々との間の相互相関値を導出することと、  
前記比較ステップで導出された最大相互相関値を決定し、かつ前記最大相互相関値が前記真の異常を前記第3テンプレートまたは前記第4テンプレートのどちらと比較することによって生成されたかを決定することと、  
前記最大相互相関値が前記真の異常を前記第3テンプレートと比較することによって生成された場合には、前記真の異常を悪性異常と分類し、前記最大相互相関値が前記真の異常を前記第4テンプレートと比較することによって生成された場合には、前記真の異常を良性異常と分類することとをさらに具備する請求項17に記載の方法。

10

【請求項25】

医用デジタル画像の異常を分類する方法において、  
予め定められた悪性異常および予め定められた良性異常にそれぞれ対応する複数の第1テンプレートおよび複数の第2テンプレートを得ることと、  
前記医用デジタル画像の前記異常を前記得られた第1および第2テンプレートと比較して、前記異常と前記得られた第1および第2テンプレートの各々との間の相互相関値を導出することと、  
前記比較ステップで導出された最大相互相関値を決定し、かつ前記最大相互相関値が前記異常を前記第1テンプレートまたは前記第2テンプレートのどちらと比較することによって生成されたかを決定することと、  
前記最大相互相関値が前記異常を前記第1テンプレートと比較することによって生成された場合には、前記異常を悪性異常と分類し、前記最大相互相関値が前記異常を前記第2テンプレートと比較することによって生成された場合には、前記異常を良性異常と分類することと  
を具備する改良された方法。

20

【請求項26】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、  
前記医用デジタル画像の前記異常よりサイズが小さい複数の第1テンプレートおよび複数の第2テンプレートを得ることを具備する請求項25に記載の方法。

30

【請求項27】

前記比較工程が、  
前記候補異常を前記得られた第1および第2テンプレートと比較して、前記候補異常と前記得られた第1および第2テンプレートの各々との間の相互相関値を導出することと、  
前記第1および第2テンプレートを前記候補異常に対してシフトさせて、前記異常と前記シフトされた第1および第2テンプレートの各々との間の相互相関値を導出することとを具備する請求項26に記載の方法。

40

【請求項28】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、  
前記得られた複数の第1テンプレートおよび複数の第2テンプレートの鏡像である追加第1および第2テンプレートを生成することを具備する請求項25に記載の方法。

【請求項29】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、  
前記得られた複数の第1テンプレートのスケーリングされたバージョンである追加第1テンプレートを生成することを具備する請求項25に記載の方法。

【請求項30】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、

50

前記得られた複数の第1テンプレートおよび複数の第2テンプレートの回転バージョンである追加第1および第2テンプレートを生成することを具備する請求項25に記載の方法。

【請求項31】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートおよび複数の第2テンプレートの鏡像である追加第1および第2テンプレートを生成することを具備する請求項26に記載の方法。

【請求項32】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートのスケーリングされたバージョンである追加第1テンプレートを生成することを具備する請求項26に記載の方法。 10

【請求項33】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートおよび複数の第2テンプレートの回転バージョンである追加第1および第2テンプレートを生成することを具備する請求項26に記載の方法。

【請求項34】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートおよび複数の第2テンプレートの鏡像である追加第1および第2テンプレートを生成することを具備する請求項27に記載の方法。 20

【請求項35】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートのスケーリングされたバージョンである追加第1テンプレートを生成することを具備する請求項27に記載の方法。

【請求項36】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートおよび複数の第2テンプレートの回転バージョンである追加第1および第2テンプレートを生成することを具備する請求項27に記載の方法。

【請求項37】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートのスケーリングされたバージョンである追加第1テンプレートを生成することを具備する請求項34に記載の方法。 30

【請求項38】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、前記得られた複数の第1テンプレートおよび複数の第2テンプレートの回転バージョンである追加第1および第2テンプレートを生成することを具備する請求項35に記載の方法。

【請求項39】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、候補第1テンプレートおよび候補第2テンプレートを得ることと、前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項25に記載の方法。 40

【請求項40】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、候補第1テンプレートおよび候補第2テンプレートを得ることと、 50

前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、

前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項26に記載の方法。

【請求項41】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、

候補第1テンプレートおよび候補第2テンプレートを得ることと、

前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、

10

前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項27に記載の方法。

【請求項42】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、

候補第1テンプレートおよび候補第2テンプレートを得ることと、

前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、

20

前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項34に記載の方法。

【請求項43】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、

候補第1テンプレートおよび候補第2テンプレートを得ることと、

前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、

30

前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項37に記載の方法。

【請求項44】

前記複数の第1テンプレートおよび複数の第2テンプレートを得る工程が、

候補第1テンプレートおよび候補第2テンプレートを得ることと、

前記候補第1テンプレートを少なくとも1つの既知の非異常構造と比較し、前記少なくとも1つの既知の非異常構造に対し予め定められた程度 of 非類似性を示した候補第1テンプレートだけを第1テンプレートとして保持することと、

40

前記候補第2テンプレートを少なくとも1つの既知の異常構造と比較し、前記少なくとも1つの既知の異常構造に対し予め定められた程度 of 非類似性を示した候補第2テンプレートだけを第2テンプレートとして保持することとを具備する請求項38に記載の方法。

【請求項45】

請求項1～44のいずれか1項に記載の方法を実現するためのシステム。

【請求項46】

請求項1～44のいずれか1項に記載のステップを実行するためのプログラムを格納するコンピュータ読み取り可能な媒体。

【発明の詳細な説明】

【0001】

(連邦政府の資金援助による研究に関する陳述)

50



本発明は部分的に、U S P H S 許諾番号 C A 6 2 6 2 5 および C A 6 4 3 7 0 (国立衛生研究所)に基づく米国政府の支援により行なわれた。米国政府は当該発明に特定の権利を有する。

**【0002】**

(発明の分野)

本発明は、医用画像における肺小結節の自動検出のためのプロセス、システム、およびコンピュータ読み取り可能な媒体に関する。

**【0003】**

本発明はまた一般的に、例えば米国特許第4, 839, 807号、第4, 841, 555号、第4, 851, 984号、第4, 875, 165号、第4, 907, 156号、第4, 918, 534号、第5, 072, 384号、第5, 133, 020号、第5, 150, 292号、第5, 224, 177号、第5, 289, 374号、第5, 319, 549号、第5, 343, 390号、第5, 359, 513号、第5, 452, 367号、第5, 463, 548号、第5, 491, 627号、第5, 537, 485号、第5, 598, 481号、第5, 622, 171号、第5, 638, 458号、第5, 657, 362号、第5, 666, 434号、第5, 673, 332号、第5, 668, 888号、第5, 740, 268号、第5, 790, 690号、第5, 832, 103号、第5, 873, 824号、第5, 881, 124号、第5, 931, 780号、第5, 974, 165号(PCT公報WO95/15537)、第5, 982, 915号、第5, 984, 870号、第5, 987, 345号、第6, 011, 862号、第6, 058, 322号、第6, 067, 373号、第6, 075, 878号、第6, 078, 680号、第6, 088, 473号、第6, 112, 112号、第6, 138, 045号、および第6, 141, 437号に加えて、米国特許出願第08/173, 935号、第08/900, 188号、第08/900, 189号、第08/979, 639号、第08/982, 282号、第09/027, 468号、第09/028, 518号、第09/092, 004号、第09/121, 719号、第09/141, 535号、第09/298, 852号、および第09/471, 088号、ならびに米国特許仮出願第60/107, 095号、第60/160, 790号、第60/176, 297号、第60/176, 304号、第60/180, 162号、第60/193, 072号、および第60/207, 401号の1つまたはそれ以上に開示されているようなデジタル画像の自動解析のためのコンピュータ化技術にも関する。これらの全ては引用文献としてここに組み込まれる。

**【0004】**

本発明は、上記の米国特許および特許出願で参照され記載されているだけでなく、添付の付属書で著者および発行年によって識別されると共に明細書中で付属書に列記されたそれぞれの参考文献に対応する括弧内の太字の数字によって相互参照される参考文献にも記載されている、様々な技術の使用を含む。上記の関連特許および出願ならびに付属書に列記する文献を含め、それらの内容全体は引用文献としてここに組み込まれる。

**【0005】**

(背景の説明)

放射線専門医は胸部放射線写真で肺小結節を陽性事例の30%も検出し損なうことがあると報告されている[1, 2]。放射線専門医によって見落とされた肺癌の多くは、振り返ってみると、実際には以前の放射線写真で見えていた[3]。したがって、シカゴ大学放射線医学部の発明者らおよび他の人々は、デジタル胸部放射線写真における肺小結節の検出で放射線専門医を支援するためにコンピュータ支援診断(CAD)方式を開発した[4~9]。先在する方式の1つの問題は、自動化方式によって生じる比較的多数の偽陽性であり、それは小結節の検出に対するCAD方式の臨床応用における主要な障害を構成している。

**【0006】**

(発明の概要)

したがって、本発明の目的は、小結節と誤って報告される偽陽性の数がそれによって低減

されるCADプロセス、システム、およびコンピュータプログラム製品を提供することである。

#### 【0007】

これおよび他の目的は、本発明に従って、医用デジタル画像における候補異常が真の異常であるかどうかを決定する新規の改善された方法、該方法を実現するシステム、および該方法を実現するプログラムステップを格納するコンピュータ読み取り可能な媒体を提供することによって達成される。該方法は、予め定められた異常および予め定められた非異常にそれぞれ対応する複数の第1テンプレートおよび複数の第2テンプレートを得ることと、候補異常を得られた第1および第2テンプレートと比較して、候補異常と得られた第1および第2テンプレートの各々との間の相互相関値を導出することと、比較工程で導出された最大相互相関値を決定し、該最大相互相関値が候補異常を第1テンプレートまたは第2テンプレートのどちらと比較することによって生成されるかを決定することと、最大相互相関値が候補異常を第1テンプレートと比較することによって生成される場合には、候補異常を真の異常であると決定し、最大相互相関値が候補異常を第2テンプレートと比較することによって生成される場合には、候補異常を非異常であると決定することを含む。悪性および良性異常に対応する追加テンプレートとの比較によって得られるさらなる相互相関値に基づいて、真の異常が悪性かそれとも良性かが同様に決定される。

10

#### 【0008】

小結節は一般的に非小結節テンプレートより小結節テンプレートによく類似しており、かつ非小結節は小結節テンプレートより非小結節テンプレートによく類似しているので、候補小結節の各々に対して小結節テンプレートおよび非小結節テンプレートにより得られる最大相互相関値は、非小結節を小結節から区別するために使用される。したがって、小結節テンプレートによる小結節の最大相互相関値は一般的に、非小結節テンプレートによるものより大きく、逆もまた同じである。したがって、本発明では、得られる最大相互相関値が決定され、次いで、最大相互相関値が小結節テンプレートとの比較に基づいて得られる場合には、候補小結節は真の小結節であると決定され、最大相互相関値が非小結節テンプレートとの比較に基づいて得られる場合には、候補小結節は偽の小結節であると決定される。

20

#### 【0009】

本発明のCADプロセスを実現する研究が行なわれ、それによって、多重テンプレートマッチング技術の使用により、非常に少数の真の陽性(2.3%)の低減と共に、胸部放射線写真の多数の偽陽性(44.3%)が除去された。加えて、多重テンプレートマッチング技術の使用により、CT画像における小結節の検出のための別のCAD方式で、同様の結果が得られた。したがって、本発明は、胸部放射線写真の小結節、乳房X線写真の塊および微小石灰化、CT画像の結節、結腸ポリープ、肝臓腫瘍、および動脈瘤に加えて、超音波および磁気共鳴画像での胸部病変を含め、医用画像における様々な病変の検出のための多くの様々なCAD方式の性能を改善するために適用可能であると考えられる。さらに、多重テンプレートマッチング技術は、良性病変と肺癌、乳癌、結腸癌、および胃癌などの悪性病変との区別のためのCAD方式の性能を改善するために、良性病変と悪性病変とを見分ける用途を有する。

30

40

#### 【0010】

本発明およびそれに伴う利点の多くについてのより完全な理解は、以下の詳細な説明を参考にして添付の図面に関連して考察したときにそれらをより深く理解するようになるので、容易に得られるであろう。

#### 【0011】

(好適な実施形態の説明)

方法および材料

胸部放射線写真に向けられた本発明の最初の開発で使用された胸部画像は、シカゴ大学病院放射線科で選択された100枚の異常な背腹方向(PA)の胸部放射線写真で構成された[9]。全部で122個の小結節を2名の放射線専門医の合意に基づいて確認し、CT

50

スキャンまたは放射線写真による追跡調査によって検証した。小結節の直径は4 mmから27 mmの範囲であった(平均13 mm)。コニカレーザフィルムスキャナによる胸部放射線写真のデジタル化によって、デジタル画像を得た。原デジタル画像は0.175 mmの画素サイズ、2000×2000のマトリックスサイズ、および10ビットのグレーレベル範囲を有する。小結節は胸部画像で比較的大きいので、デジタル画像のマトリックスサイズは4分の1に縮小した。したがって、最終画像は0.7 mmの画素サイズおよび500×500のマトリックスサイズを有した。上記の主データベースに加えて、補助データベース、日本標準デジタル画像データベース[10]を、小結節テンプレートの作成および多重テンプレートマッチング技術の訓練のためにのみ使用し、多重テンプレートマッチング技術による本CAD方式の検証には使用しなかった。補助データベースは、日本放射線写真技術協会によって開発されたものであり、後で本書で説明する。

10

#### 【0012】

先在するCAD方式により、最初に我々のデータベース内の100枚の胸部画像から116個の小結節および4875個の偽陽性を検出した。次いで規則に基づく分類技術を適用して偽陽性の数を低減し[9]、それにより88個の小結節および377個の偽陽性を得た。多重テンプレートマッチング技術のために、最初に先在CAD方式によって検出された116個の小結節および4875個の偽陽性を、それぞれ初期小結節テンプレートおよび初期非小結節テンプレートとして使用した。しかし、最終的に本CAD方式によって報告された88個の小結節および377個の偽陽性を、訓練および最終結果の試験のために使用した。多重テンプレートマッチング技術の優れた性能が達成されるように、8枚の非小結節状小結節テンプレートおよび1798枚の小結節状非小結節テンプレートをそれぞれ小結節テンプレートセットおよび非小結節テンプレートセットから系統的に除去した。また、試験候補セットのうち377の偽陽性が先在する規則に基づく試験で残り、したがって小結節に似ている「難しい」偽陽性とみなされることも注目される。研究では、多重テンプレートマッチング技術を使用して、これらの難しい偽陽性の数をかなり低減した。

20

#### 【0013】

規則に基づく試験に基づく先在CAD方式の概要

まず第一に、PA胸部画像の肺領域をエッジ、肺上部、および横隔膜の使用によって区分し[11、12]、その後のCAD方式の処理のために使用した。次いで、複雑な背景構造を軽減することができ、かつしたがって小結節をより際立たせることができるように、(整合フィルタの使用による)小結節強調画像から(平滑化フィルタの使用による)小結節抑制画像を差し引くことによって差分画像を得た。差分画像から初期小結節候補を検出するために、差分画像を様々な閾レベルにより二値化することによって、複数の2値画像を得た。2値画像の各々で、成分ラベル付け技術を使用して、各々の孤立した「島」を識別し[4]、各島に対し2つの特徴、有効直径および真円度を決定した。次いで、その有効直径および真円度がそれぞれ6.5 mmおよび0.65に等しいかそれより大きければ、島を初期小結節候補であるとみなした。100枚の異常胸部画像のデータベースに対し、全部で116個(122個のうち)の小結節および4875個の偽陽性を初期小結節候補として識別した。

30

#### 【0014】

次に、差分画像および原画像両方の初期小結節候補の位置に、2つの画像の各々の背景から小結節候補を正確に区分するために、領域成長技術[5、6]を適用した。次いで、有効直径、コントラスト、真円度、不正則度、エッジ勾配、真円度勾配、不正則度勾配、および直径勾配など様々な特徴を各々の成長領域から決定し、規則に基づく分類技術によって小結節と偽陽性とはを見分けるために使用した。このステップ後に、大部分の小結節、116個のうち88個(75.9%)が保持され、大部分の偽陽性、4875個のうち4498個(92.3%)が除去され、したがって377個の偽陽性が残された。

40

#### 【0015】

偽陽性の低減のための多重テンプレートマッチング技術の全方式

図1は、本発明の多重テンプレートマッチングプロセスの全方式を示す。まず第一に、ス

50

ステップ10で、デジタル胸像が得られ、ステップ20で小結節検出のためのCAD方式が適用され、ステップ30で候補小結節がそれによって得られる。ステップ40および50で、小結節テンプレートセットおよび非小結節テンプレートセットが得られ、ステップ60で、候補小結節の各々に対して、相互相関技術が使用され、小結節テンプレートセットおよび非小結節テンプレートセットとの相互相関値が計算される。最後に、小結節テンプレートセットおよび非小結節テンプレートセットで得られた2つの最大相互相関値が決定され、ステップ70で偽陽性（非小結節）を除去するために使用される。その結果、ステップ80で、残りの候補小結節が真の小結節として識別される。最大相互相関値が候補小結節と非小結節テンプレートの比較によって得られる場合には、候補小結節は偽陽性と決定される。他方、最大相互相関値が候補小結節と小結節テンプレートの比較によって得られる場合には、候補小結節は真の小結節と決定される。

10

**【0016】**

多重テンプレートマッチング技術を適用するために、2つの多重テンプレートセット、すなわち多数の小結節テンプレートを持つ1つ、および多数の非小結節テンプレートを持つもう1つを形成した。この研究では、先在CAD方式によって報告された初期小結節候補（116個の小結節および4875個の非小結節）を初期テンプレートとして使用した。それらは初期小結節候補の位置を中心とする $36 \times 36$ 画素の小さい関心領域（ROI）であった。

**【0017】**

図2は、初期小結節テンプレートから小結節テンプレートセットを作成するための基本方式を示す。ステップ41で、初期小結節テンプレートを得、次いでステップ42で、テンプレートの数が倍になるように、各々の小結節テンプレートを最初に左右反転してミラーテンプレートを生成した。次いで、ステップ43で、小結節テンプレートの各々をスケールリング（縮小または拡大）し、ステップ44では回転もして、小結節テンプレートの数をさらに増加した。最後に、ステップ45で、8枚の非小結節状小結節テンプレート（肋骨と重なる非常に微妙な小結節などの異型小結節）を慎重に不適切な小結節テンプレートとして識別し、したがって小結節テンプレートセットから除外して、ステップ46で小結節のテンプレートセットを導出した。

20

**【0018】**

図3は、非小結節テンプレートセットを作成するための同様の方式を示す。ステップ51で得られた各々の非小結節テンプレートをステップ52で最初に左右反転し、ステップ53では回転もして、非小結節テンプレートの数を増加した。次いでステップ54で、小結節と同様の外観を有する多く（1789枚）の非小結節テンプレートが、後で本書で示す訓練セットを使用することによって、系統的に非小結節テンプレートセットから除去される。これらの小結節状非小結節テンプレートの存在は、多重テンプレートマッチング技術の性能をかなり劣化させ得るからである。比較的多数の非小結節テンプレートが得られたので、かつスケールリングの使用によって得られたテンプレートは幾分不明瞭であるので、非小結節テンプレートはスケールリングしなかった。非小結節状小結節テンプレートおよび小結節状非小結節テンプレートの除去後、小結節テンプレートセットで得られた108枚の小結節テンプレートに加えて、非小結節テンプレートセットで3077枚の非小結節テンプレートがステップ55で得られた。

30

40

**【0019】**

図4および5は、小結節テンプレートおよび非小結節テンプレートの典型例をそれぞれ示す。図4および5で、小結節テンプレートおよび非小結節テンプレートの外観が全く異なり、したがって小結節と非小結節を見分けるための基礎であることが明瞭である。

**【0020】**

先在CAD方式によって最終的に報告された88個の小結節を次いで、各々44個の無作為に選択された小結節を有する2つのセットに分割した。訓練セットと呼ばれる1セットは、本書で後述する通り、非小結節テンプレートセットから小結節状非小結節テンプレートを系統的に除去するために使用した。試験セットと呼ばれる他のセットの44個の小結

50

節と377の非小結節（偽陽性）は、本発明の多重テンプレートマッチング技術の性能を検証するために使用した。テンプレートと同様に、試験候補は、試験セットの小結節候補の位置を中心として40×40画素の小さいROIであった。

#### 【0021】

テンプレートのマトリックスサイズの決定およびテンプレートのスケーリングおよび回転など多数のパラメータの効果の調査のための計算時間を低減するために、上記のデータセットに加えて、小さいテンプレートセットおよび小さい試験セットも使用した。小さいテンプレートセットは小結節テンプレートセットの108枚の小結節テンプレート全部、および4875個の非小結節から手動で選択された178枚の「典型的」非小結節テンプレートを含んでいた。小さい試験セットは、88個の小結節および377個の非小結節からそれぞれ無作為に選択された30個の小結節および30個の非小結節を含んでいた。

10

#### 【0022】

試験候補と小結節および非小結節のテンプレートセットとの間の最大相互相関値の決定テンプレートと試験候補との間の相互相関値を決定するために、総当たり法、すなわち総当たり（leave-one-out）試験法を使用した。すなわち、試験候補がテンプレートのセットに含まれていれば、対応するテンプレートは試験候補との相互相関値の決定には使用しない。テンプレートと試験候補の間の相互相関値を実際に計算する前に、テンプレートおよび試験候補に含まれる背景傾向の補正のために前処理ステップを利用した。テンプレート（または試験候補）の背景傾向を2次元（表面）線形関数によって表わし、最小2乗法によって線形関数の係数を決定した。次いで、テンプレート（または試験候補）の原画像から推定表面関数を減算して背景傾向補正画像を出した。

20

#### 【0023】

図6および7は、テンプレートの原画像および推定背景をそれぞれ示す。原画像から推定背景を減算することによって得た背景傾向補正画像を図4に示す。

#### 【0024】

試験候補のマトリックスサイズ（40×40画素）は、水平および垂直両方向でテンプレートのそれ（36×36画素）より4画素大きい。テンプレートAと試験候補Bとの間の相互相関値を決定するために、テンプレートAを水平および垂直両方向に4画素の最大シフト値分移動し、次式によって各シフト値（i,j）における相互相関値C<sub>ij</sub>を計算した。

30

#### 【数1】

$$C_{ij}^2 = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \frac{\{A(m,n) - \bar{A}\} \{B_{ij}(m,n) - \bar{B}\}}{\sigma_A \sigma_B}, \quad i,j \in \{0,1,2,3,4\},$$

ここで $\bar{A}$ および $\sigma_A^2$ はそれぞれテンプレートAの画素値の平均および変分であり、 $\bar{B}$ および $\sigma_B^2$ はそれぞれシフト値（i,j）における部分試験候補Bの部分領域B<sub>ij</sub>の画素値の平均および変分である。領域AおよびB<sub>ij</sub>の画素値の平均および変分は、次の方程式によって定義される。

40

#### 【0025】

#### 【数2】

$$\bar{A} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N A(m,n),$$

$$\sigma_B^2 = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \{B_{i,j}(m,n) - \bar{B}\}^2.$$

10

$$\bar{B} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N B_{i,j}(m,n),$$

$$\sigma_A^2 = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \{A(m,n) - \bar{A}\}^2,$$

## 【0026】

20

次いで、全シフト値中の最大相互相関値をテンプレートと試験候補との間の相互相関値と決定した。次に、全ての小結節テンプレートと試験候補の相互相関値を計算し、最大相互相関値を決定し、小結節との類似の程度を示す試験候補の独特の特徴として使用した。同様に、非小結節テンプレートと試験候補の最大相互相関値を決定し、非小結節との類似の程度を示す試験候補の他の独特の特徴として使用した。次いで、非小結節テンプレートとの最大相互相関値が小結節テンプレートとのそれより大きい場合、偽陽性と小結節を見分けるために試験セットにおける候補のこれらの2つの特徴を使用した。

## 【0027】

小結節テンプレートおよび非小結節テンプレートの作成

多重テンプレートマッチング技術の成功のための重要な要素は、小結節および非小結節に利用可能なテンプレートの数である。この研究では、小結節検出小結節のための我々のCAD方式によって最初に検出された116の小結節から選択された108枚だけを小結節テンプレートとして使用した。小結節を含まない胸部放射線写真のどのROIも理論的に非小結節テンプレートとみなすことができるが、先在するCAD方式によって最初に検出された4875枚の偽陽性から、3077枚の偽陽性が「典型的な」非小結節とみなされたので、それらを非小結節テンプレートとして選択した。次いで次の3つの方法を利用して、この研究のテンプレートの数を増加した。

(1) テンプレートを左右反転してミラーテンプレートを作成する。

(2) 0.6、0.8、および1.2の3つの異なる倍率にテンプレートをスケーリングする。

40

(3) 2つの異なる角度、-10度および+10度テンプレートを回転する。

原画像から直接得たROIを原画テンプレートと呼び、原画テンプレートの左右反転、スケーリング、または回転によって得られたものを導出テンプレートと呼んだ。追加倍率によるスケーリングによって、かつ追加角度の回転によっても、追加導出テンプレートを作成できることを注目されたい。

## 【0028】

左右反転は、右肺のテンプレートを左肺に役立てることができ、その逆も真であり、それはテンプレートの数を2倍にする。この研究ではスケーリングおよび回転によりテンプレートの数をそれぞれ4倍および3倍に増加し、したがってテンプレートの数はこれらの方法の組合せによって24倍に増加することができる。

50

## 【0029】

図8は、小さい試験セットにおける60個の候補について、小さいテンプレートセットにおける108枚の小結節テンプレートおよび178枚の非小結節テンプレートとの最大相互相関値間の関係を示す。前述した通り、テンプレートは典型的な小結節および非小結節から慎重に選択した。図8で、かなりの重複があるが、小結節は小結節テンプレートとより大きい最大相関値を持つ傾向がある一方、非小結節は非小結節テンプレートとそうなり、すなわち小結節は45度の線より上にあり、非小結節は45度の線より下にある傾向がある。この一般的傾向は、小結節と非小結節を見分ける上での多重テンプレートマッチング技術の有用性を示す。しかし、多くの試験候補が小結節または非小結節テンプレートのいずれとも比較的低い相互相関値を持ち、それは、小結節または非小結節の試験セットに、これらの試験候補に非常に類似した十分なテンプレートが無かったことを示す。

10

## 【0030】

図9は、(a)ミラー小結節テンプレートのみ、(b)ミラー非小結節テンプレートのみ、および(c)小結節および非小結節の両方のミラーテンプレートに基づいて追加テンプレートを組み込んだ場合、60個の試験候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示す。図8および図9(a)を比較すると、図9(a)では大部分の試験候補が上方に移動し、したがって、小結節のミラーテンプレートを使用することにより、これらの試験候補に対し小結節テンプレートとの最大相互相関値が増加したことが明らかである。加えて、小結節は図9(a)の対角線上またはそれより上に位置しており、したがって対角線より下の幾つかの非小結節は小結節から明瞭に見分けることができる。同様に、図8と比較して、図9(b)では大部分の試験候補が右側に移動しており、したがって、非小結節のミラーテンプレートの使用により、非小結節テンプレートとの最大相互相関値が増加したことが観察された。図9(c)では、試験候補が上方および右側に移動し、ミラーテンプレートの使用によって、小結節および非小結節テンプレートの両方との最大相互相関値が増加したことが明らかである。

20

## 【0031】

研究には非常に限られた数の小結節しか無かったので、スケーリングは場合によってはテンプレートを不明瞭にするが、それでもそれを小結節テンプレートに適用した。図10は、小結節テンプレートおよび非小結節テンプレートと共にスケーリングされた小結節テンプレートとの最大相互相関値間の関係を示す。図10の結果を図8のスケーリングされた小結節テンプレート無しの結果と比較すると、スケーリングされた小結節テンプレートの追加により、結果的に試験候補に対し小結節テンプレートの最大相互相関値が改善され、また大部分の小結節が対角線より上にあるので、小結節と非小結節の区別も改善されることが明らかである。

30

## 【0032】

他方、スケーリングと異なり、回転は小結節および非小結節テンプレートの全てに適用した。図11は、(a)小結節テンプレートのみ、(b)非小結節テンプレートのみ、および(c)小結節および非小結節テンプレート両方の回転によって追加テンプレートを形成した場合の小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示す。回転テンプレート無しの結果は図8に示す。再び、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値が、回転テンプレートの追加によって大きくなった。

40

## 【0033】

多重テンプレートマッチング技術の別のパラメータは、テンプレートおよび試験候補のマトリックスサイズの適切な選択である。この研究では、 $24 \times 24$ 画素から $48 \times 48$ 画素までの範囲の様々なマトリックスサイズの総合性能に対する効果を検討した。テンプレートおよび試験候補のマトリックスサイズがそれぞれ $36 \times 35$ 画素(約 $25 \times 25 \text{ mm}^2$ )および $40 \times 40$ 画素( $28 \times 28 \text{ mm}^2$ )のときに、最良の結果が得られることが分かった。

## 【0034】

50

図12(a)、(b)、および(c)は、テンプレートのマトリックスサイズがそれぞれ $24 \times 24$ 、 $36 \times 36$ 、および $48 \times 48$ 画素に等しいときに、60個の候補についての小結節テンプレートおよび非小結節テンプレートとの最大相互相関値の関係を示す。使用したテンプレートは、108の小結節および178の非小結節と共に、それらのミラーテンプレートを含んでいた。最大相互相関値は一般的に大きかったが、 $24 \times 24$ 画素のマトリックスサイズでは、小結節と非小結節との間の相関値の重複も大きかった。それは小結節と非小結節の分離には役立たない。他方、マトリックスサイズが $48 \times 48$ 画素のときには、最大相互相関値が通常低く、信頼できる特徴として使用することができなかった。図12から、30個の小結節と30個の非小結節との間の最良の分離は、テンプレートのマトリックスサイズが $36 \times 36$ 画素であるときに達成されたことが明らかである。したがって、この研究で使用したテンプレートのマトリックスサイズは $36 \times 36$ 画素と決定された。

10

#### 【0035】

非小結節テンプレートセット中の小結節状非小結節の除去

全ての非小結節テンプレートが多重テンプレートマッチング技術の性能の改善に有用な貢献をすることができるわけではないことに、注目することが重要である。実際、多くの非小結節テンプレートは多重テンプレートマッチング技術の性能を損なうのである。図9(c)から分かるように、対角線の下に位置する4個の小結節があり、それはこれらの小結節が、全ての小結節テンプレートより、使用した幾つかの非小結節テンプレートによく類似していることを暗示する。図13は、小結節と非小結節テンプレートとの最大相関値を増加することによって、幾つかの小結節を対角線の下に引き下げるそのような非小結節テンプレートの例を示す。これらの非小結節テンプレートは外見が一般的に小結節に似ていることが分かった。したがって、多重テンプレートマッチング技術の優れた性能を達成するために、これらの小結節状非小結節テンプレートを除去することが望ましい。訓練セットおよび補助データベースの小結節を使用して、このタスクを達成した。

20

#### 【0036】

前述した主データベースに加えて、より多くの小結節テンプレートの作成、および多重テンプレートマッチング技術の訓練のためにだけ、別の補助データベースをも使用したが、補助データベースの胸部画像の特性は主データベースのそれとは全く異なるので、多重テンプレートマッチング技術による本CAD方式の検証には使用しなかった。補助データベースは、日本放射線写真技術協会によって開発された日本標準デジタル画像データベース内の合計154件の小結節症例から選択された、孤立肺小結節を持つ128枚の胸部画像を含んでいた[10]。極めて微妙な小結節または明瞭な小結節にそれぞれ対応する1または5いずれかの微妙さ格付けスコア(subtlety rating score)を持つ小結節を各々含む、26件の小結節の症例が日本標準データベースから除去された。原胸部画像は、 $0.175\text{ mm}$ の画素サイズ、 $2048 \times 2048$ のマトリックスサイズ、および12ビットのグレーレベルでデジタル化した。この研究では、主データベースの胸部画像と一致させるために、原画像データを4分の1にサブサンプリングすることによって $512 \times 512$ に縮小し、グレーレベルの数を10ビットに低減した。補助データベースの128個の小結節を主データベースの108個の小結節と共に、この後、多重テンプレートマッチング技術の検証のために小結節テンプレートセットとして使用した。128個の小結節は、多重テンプレートマッチング技術の訓練のため、すなわち小結節状非小結節テンプレートの除去のためにも使用した。

30

40

#### 【0037】

多重テンプレートマッチング技術の訓練の第1ステップとして、128個の結節の各々について、小結節と20個の最大相互相関値を出した20枚の非小結節テンプレートをここでは小結節状非小結節とみなし、初期非小結節テンプレートセットの4,875枚の原画テンプレートから除去した。こうして1,338枚の原画非小結節テンプレートを除去し、3,537枚の原画非小結節テンプレートを非小結節テンプレートセットに残した。同様に、3,537枚の原画テンプレートから小結節状非小結節をさらに除去するために、

50



訓練セットの44個の小結節を使用した。こうして、460枚の原画非小結節テンプレートを再び除去し、最終的に3,077枚の原画テンプレートを非小結節テンプレートセットに残した。補助データベースの128個の小結節および訓練セットの44個の小結節と一緒に使用することによって、合計1,798枚のテンプレートを初期非小結節テンプレートセットから除去した。テンプレートセット中の小結節状非小結節の除去のために、経験的に20個の最大相互相関値を閾値として使用した。しかし、利用可能なテンプレートの数および特性によって、異なる数の最大相互相関値を使用することが可能である。

#### 【0038】

我々が小結節状非小結節の除去によって多重テンプレートマッチング技術の性能をいかに改善したかを実証するために、図14に、小結節状非小結節テンプレートを除去する前（xで示す）および後（丸で示す）の試験セットの44個の小結節に対する小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示す。図14で使

小結節テンプレートセットは5,664（24×236）枚のテンプレートから成り、それは主データベースの108個の小結節および補助データベースの128個の小結節、ならびのそれらのミラーテンプレート、スケールリングされたテンプレート、および回転テンプレートを含んでいた。非小結節テンプレートセットは、小結節状非小結節テンプレートの除去前および除去後にそれぞれ29,250（6×4,875）枚のテンプレートおよび18,462（6×3,077）枚のテンプレートを含み、それは原画非小結節テンプレート、それらのミラーテンプレート、および回転テンプレートを含んでいた。図14で、小結節状非小結節テンプレートの除去によって大部分の小結節が左側にかなり移動したこと、および1つを除いて全ての小結節が、予測できる通り、対角線の上に位置することが明瞭である。

#### 【0039】

小結節状非小結節テンプレートの除去は、試験セット中の非小結節にも影響を及ぼすことに注目されたい。図15は、非小結節テンプレートの除去前（xで示す）および除去後（点で示す）の試験セットの377個の非小結節の半分（189個）について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示す。小結節状非小結節テンプレートの除去前および除去後に、小結節テンプレートセットには5,664枚のテンプレート、非小結節テンプレートセットには29,250枚および18,462枚のテンプレートがそれぞれあったことに注意されたい。より分かりやすい表示のために、非小結節の半分だけが図15には示されている。図15で使

用した小結節テンプレートセットおよび非小結節テンプレートセットは、図14で使

用したものと同一である。図15では、377個の非小結節のうちの幾つかが図14の小結節の場合と同様に左側に移動したが、多くの非小結節は、小結節状非小結節テンプレートの除去後も依然として、対角線の下に維持されることが明らかである。この発見は、小結節状非小結節テンプレートの除去後、ほとんど全ての小結節が対角線より上に位置するので、図15の対角線より下の偽陽性（非小結節）を図14の小結節と区別できることを暗に示す。

#### 【0040】

小結節状非小結節の除去後に、5,664枚の小結節テンプレートおよび18,462枚の非小結節テンプレートにより、試験セットの44個の小結節および377個の非小結節に基づき、検証試験を実行した。44個の小結節および377個の非小結節は、多重テンプレートマッチング技術の訓練、すなわちテンプレートセット中の小結節状非小結節または非小結節状小結節いずれかの除去には使用されておらず、図15に示した377個の非小結節は小結節状非小結節テンプレートを除去する効果を実証するために示したものであることに注目されたい。図16は、試験セットの44個の小結節および377の非小結節について、5,664枚の小結節テンプレートおよび18,462枚の非小結節テンプレートとの最大相関値間の関係を示す。図16では、2つの最大相互相関値に基づいて、小結節と偽陽性（非小結節）との間で顕著な区別ができることが明らかである。例えば対角線を閾値として使用した場合、すなわち対角線の上に位置する候補を小結節として受け入れる場合、377個の非小結節から167個の偽陽性（44.3%）を除去することがで

き、真の小結節の除去はわずか1個(2.3%)である。377個の偽陽性の多くが外見的に小結節に類似しており、したがって「困難な」偽陽性とみなされるので、これは実際に、既存のCAD方式のかなりの改善を構成する。

#### 【0041】

多重テンプレートマッチング技術は、多数のテンプレートとの相互相関値が計算されるので、比較的大量のコンピュータ時間を必要とする。例えば、最終検証試験で24,126(5,664+18,462)枚のテンプレートを使用した場合、インテル・ペンティアムⅠ1173MHzCPUおよびLinuxオペレーティングシステムを持つパーソナルコンピュータで、各々の試験候補に対し相互相関値を計算するのに約85秒かかる。将来、発明者らは、より高い性能および信頼性を達成するために、そのデータベースを大幅に拡張する計画であり、これが今度は、より大きいテンプレートセットおよび相互相関値の計算に要するより長いコンピュータ時間を暗に示す。しかし、この技術はCAD方式の最終ステップでのみ適用され、それは平均して1枚の胸部画像につき3個または4個の小結節候補を報告するだけであるので、これは多重テンプレートマッチング技術にとって重大な問題ではない。さらに、コンピュータは今も今後もどんどん高速化し、したがって、多重テンプレートマッチング技術の大量の計算によって生じる問題を相当軽減するであろう。現在、多重テンプレートマッチング技術を使用することによって幾つかの偽陽性を除去するために、各々の胸部画像を処理するのに約5~6分かかる。

#### 【0042】

コンピュータ断層撮影法への適用

多重テンプレートマッチング技術が、コンピュータ断層撮影法(CT)での小結節検出のためのCAD方式における小結節と偽陽性を区別する潜在的可能性を持つことを実証するために、合計237個の小結節を含む44のCTスキャンのデータベースに多重テンプレートマッチング技術を適用する予備研究を実施した。10mmのコリメーションおよび10mmの再構成間隔でCTスキャンを得た。各CTスライスには512×512画素のマトリックスサイズおよび10ビットのグレースケール範囲を有する。データベースの画素サイズが0.566から0.781mmの範囲となるように、視野は各患者毎に検査中に最適化した。

#### 【0043】

シカゴ大学放射線医学部におけるCT小結節検出のための既存のCADプロセス[13]は、二値化技術およびローリングボールアルゴリズムを使用することによって、最初に各スライスの背景から肺を区分した。次いで、初期小結節候補を検出するために、区分された肺領域に多重グレースケール二値化技術を適用した。初期小結節候補の各々に対し、6つの幾何学的特徴(体積、真球度、等積球体の半径、最大コンパクトさ、最大真円度、および最大偏心率)および3つのグレースケールの特徴(平均グレースケール、グレースケールの標準偏差、および小結節候補が最初に検出されたグレースケール閾値)を決定し、偽陽性と小結節を区別するために使用した。このCAD方式により、208個(87.8%)の小結節および4923個の偽陽性(1スライスにつき約3個の偽陽性)が小結節候補として報告された。

#### 【0044】

多重テンプレートマッチング技術をCT小結節検出方式に適用するための基礎を構築するために、4923個の偽陽性から無作為に選択した208個の小結節および200個の偽陽性を最初に選択した。この研究では、これらの候補から得られるテンプレートが完全に肺の内部に含まれるように、肺領域の境界から遠い小結節および非小結節だけを手動で選択した。こうして208個の小結節および200個の非小結節からそれぞれ60個の小結節および60個の非小結節を得、テンプレートおよび試験候補の両方として使用した。試験候補および対応する導出テンプレートが、試験に使用するテンプレートセットに含まれないように、総当たり試験法を使用した。上述した多重テンプレートマッチング技術により、図17に示すように、1,440(24×60)枚の小結節テンプレートおよび1,440(24×60)枚の非小結節テンプレートを使用することによって、60個の小結

10

20

30

40

50

節（丸）および60個の非小結節（点）に対して最大相互相関値を得た。ミラーテンプレート、スケールされたテンプレートおよび回転テンプレートを使用することによって、60個の小結節および60個の非小結節からそれぞれ1,440枚の小結節テンプレートおよび1,440枚の非小結節テンプレートを得た。破線は小結節と非小結節の区別のための閾値として使用した。

#### 【0045】

図17で、小結節が対角線より上に位置する傾向がある一方、非小結節は対角線より下に位置する傾向があることが明らかであり、これは小結節が一般的に小結節テンプレートによく類似しており、非小結節が一般的に非小結節テンプレートにより類似していることを示す。図17では、多くの小結節が小結節テンプレートと非常に大きい相互相関値を持ち、したがってそれらは対角線より上および上右隅付近に分布することに注目することが重要であり、これはこれらの小結節の各々がテンプレートセット中に少なくとも1つの類似した小結節テンプレートを有することを示す。したがって、CTスキャンの方が背景構造がより単純なので、CTスキャンでは胸部放射線写真の場合より2つの同様の小結節を見つける確率が大きいと考えられる。破線を閾値として使用した場合、すなわち破線より上の候補を小結節と認め、破線より下の候補を非小結節と認めた場合、29個（48.3%）の偽陽性（非小結節）をわずか1個（1.7%）の小結節の低減と共に除去することができる。この結果は、多重テンプレートマッチング技術が、CT小結節検出のためのCAD方式、および乳房X線写真における塊および微小石灰化、超音波および磁気共鳴画像における胸部病変、腹部CT画像における結腸ポリープおよび肝臓腫瘍、ならびに脳CT画像における動脈瘤など、多くの異なる種類の病変を検出するための多くのCAD方式における偽陽性の数を相当低減する潜在的な可能性を有することを示す。加えて、多重テンプレートマッチング技術は、良性病変と肺癌、乳癌、結腸癌、および胃癌など多くの胃癌による悪性病変とを区別するためにCADの性能を改善するために、良性小結節を悪性小結節と区別するのに使用することができる。

#### 【0046】

良性小結節への適用

デジタル胸部画像の小結節を区別するためのCAD方式で、多重テンプレートマッチング技術を使用して良性小結節と悪性小結節を区別できることを示すための予備研究も行なった。23個の悪性小結節および33個の良性小結節を含む、56枚の胸部画像のデータベースを使用した。各画像は2048×2048画素のマトリックスサイズ、0.175mmの画素サイズ、および10ビットのグレースケール範囲を持っていた。各々の小結節の位置およびサイズを3名の放射線専門医によって識別し、位置およびサイズの平均値を計算し、次に述べる通り、胸部画像における悪性小結節と良性小結節の区別のために使用した。

#### 【0047】

多重テンプレートマッチング技術を適用することによって良性小結節と悪性小結節を区別するために、10個の「典型的」悪性小結節および10個の「典型的」良性小結節を手動で選択した。おおよその小結節サイズが分かるので、小結節の面積がテンプレートの面積の約半分になるように、最初に可変マトリックスサイズで小結節の位置で原画テンプレートを得た。次いで、画像スケール技術を用いることによって、全てのテンプレートのマトリックスサイズを40×40画素に正規化（縮小または拡大）した。各々の小結節に対し、対応するテンプレートと同一中心位置を持ちマトリックスサイズが48×48画素である対応する試験候補のスケールングのために、同一倍率を適用した。

#### 【0048】

次いで、放射線専門医による推定小結節サイズの誤差を補正するために、各テンプレートを左右反転し、-10度および+10度の2つの角度だけ回転し、0.9および1.1の2つの倍率でスケールした。テンプレートの数を増加するために3つの技術を組み合わせることによって、我々は180（18×10）個の悪性小結節のテンプレートおよび180（18×10）個の良性小結節のテンプレートを得た。全てのテンプレートおよび

試験候補について、それらの間の相互相関値を決定する前に、背景傾向を補正した。試験候補および対応する導出テンプレートが試験に使用するテンプレートセットに含まれないように、総当たり法を使用した。

#### 【0049】

図18は、180(18×10)枚の悪性小結節テンプレートおよび180(18×10)枚の良性小結節テンプレートを使用することにより、多重テンプレートマッチング技術で得た10個の悪性小結節(丸)および10個の良性小結節(点)の最大相互相関値を示す。図18で、悪性小結節は対角線より上に位置する傾向がある一方、良性小結節は対角線より下に位置する傾向があることが明らかであり、これは悪性小結節が一般的に悪性小結節テンプレートにより類似し、良性小結節が一般的に良性小結節テンプレートにより類似することを示す。対角線を閾値として使用すると、すなわち、対角線より上の候補を悪性小結節として受け入れ、対角線より下の候補を良性小結節として受け入れると、全ての悪性小結節を残しながら、8個(80%)の良性小結節を除去することが可能である。この結果は、良性病変と肺癌、乳癌、結腸癌、および胃癌など多くの癌による悪性病変の区別のためのCAD方式の性能を改善するために、多重テンプレートマッチング技術が良性小結節と悪性小結節を区別する能力を有することを示している。

10

#### 【0050】

##### コンピュータプログラム製品

本説明で記載した機構およびプロセスは、関連技術分野の熟練者には理解される通り、本明細書の教示に従ってプログラムした従来の汎用マイクロプロセッサまたはコンピュータを使用して実現することができる。適切なソフトウェアコーディングは、同じく関連技術分野の熟練者には理解される通り、本開示の教示に基づいて熟練プログラマが容易に作成することができる。しかし、当業者には容易に明らかになる通り、本発明は、特定用途向集積回路の作製によって、あるいは従来の部品回路の適切なネットワークを相互接続することによって実現することもできる。

20

#### 【0051】

したがって本願は、記憶媒体でホスト化(hosted)することができ、かつ汎用マイクロプロセッサまたはコンピュータをプログラムして本発明に係るプロセスを実行するために使用することのできる命令を含む、コンピュータ利用製品をも含む。この記憶媒体は、フロッピーディスク、光ディスク、CD-ROM、磁気光ディスク、ROM、RAM、EPROM、EEPROM、フラッシュメモリ、磁気または光カードをはじめとする任意の種類のディスク、あるいは電子命令を格納するのに適した任意の種類の媒体を含むことができるが、これらに限定されない。

30

#### 【0052】

汎用マイクロプロセッサまたはコンピュータのプログラミングは、画像収集装置(図示せず)から得た画像をデジタル化しかつ格納するためのソフトウェアモジュールを含むことができる。代替的に、本発明はまた、画像アーカイブ通信システム(PACS)など他の手段によって得た画像から、またはデジタル画像データを生成する撮像装置から直接、導出されるデジタルデータを処理するために実現することもできる。言い換えると、処理されるデジタル画像は、デジタル形式ですでに存在しており、本発明を実施する際にデジタル形式に変換する必要のないものであるかもしれない。

40

#### 【0053】

上記の教示に照らして本発明の多数の変形および変化が可能である。したがって、請求の範囲内で本発明はここで特に記載した以外の仕方で実施できることを理解されたい。

#### 【0054】

##### 参考文献のリスト

1. J. V. Forrest and P. J. Friedman, "Radiologic errors in patients with lung cancers," West J. Med. 134, 485-490, (1981).
2. D. P. Naidich, E. A. Zerhouni, and S. S.

50

- Slegelman, Computer tomography of thorax (Raven, New York, 1984), pp. 171-199.
3. J. R. Muhm, R. S. Miller, R. S. Fontana, et al, "Lung Cancer detected during a screening program using four-month chest radiographs," *Radiology* 148, 609-615 (1983).
  4. M. L. Giger, K. Doi, and H. MacMahon, "Image feature analysis and computer-aided diagnosis in digital radiography. 3. Automated detection of nodules in peripheral lung fields," *Med. Phys.* 15, 158-166 (1988). 10
  5. M. L. Giger, K. Doi, H. MacMahon, C. E. Metz, and F. F. Yin, "Pulmonary nodules: computer aided detection in digital chest image," *RadioGraphics* 10, 41-51 (1990).
  6. T. Matsumoto, H. Yoshimura, K. Doi, M. L. Giger, A. Kano, H. MacMahon, K. Abe, and S. M. Montner, "Image featyre analysis of false-positive diagnoses produced by automated detection of lung nodules," *Invest. Radiol* 27, 587-597 (1992). 20
  7. Y. C. Wu, K. Doi, M. L. Giger, C. E. Metz, and W. Zhang, "Reduction of false positives in computerized detection of lung nodules in chest radiographs using artificial neural networks, discriminant analysis and a rule-based scheme," *J. Digital Imag.* 7, 196-207 (1994).
  8. T. Kobayashi, X. W. Xu, H. MacMahon, C. E. Metz, and K. Doi, "Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs." *Radiology* 199, 843-848 (1996). 30
  9. X. W. Xu, K. Doi, T. Kobayashi, H. MacMahon, and M. L. Giger, "Development of an improved CAD scheme for automated detection of lunf nodules in digital chest images," *Med. Phys.* 24, 1395-1403 (1997).
  10. J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Koderu, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologigts' detection of pulmonary nodule s," *AJR* 147, 71-74 (2000). 40
  11. X. W. Xu, K. Doi, "Image feature analysis for computer-aided diagnosis: Accurate determination of ribcage boundary in chest radiographs," *Med. Phys.* 22, 617-626 (1995) 50

12. X. W. Xu, K. Doi, "Image feature analysis for computer-aided diagnosis: Detection of right and left hemidiaphragm edges and delineation of lung field in chest radiographs," Med. Phys. 23, 1613-1624 (1996).
13. S. G. Armato III, M. L. Giger, C. J. Moran, J. T. Blackburn, K. Doi, and H. MacMahon, "Computerized detection of pulmonary nodules on CT scans," Radiographics 19, 1303-1311, (1999).

【図面の簡単な説明】

【図1】本発明に係る多重テンプレートマッチングプロセス全体を示すフローチャート。

【図2】小結節テンプレートセットの作成のための基本的プロセスを示すフローチャート。

【図3】非小結節テンプレートセットの作成のための基本的プロセスを示すフローチャート。

【図4】背景傾向を補正した小結節テンプレートの複数の例を示す図。

【図5】背景傾向を補正した非小結節テンプレートの複数の例を示す図。

【図6】小結節の原画像の複数の例を示す図。

【図7】図6の原画像に線形関数を当てはめることによって得られた推定背景画像の複数の例を示す図。

【図8】60の無作為に選択された候補について、108枚の小結節テンプレートおよび178枚の非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図9A】ミラー小結節テンプレート（216枚の小結節および178枚の非小結節テンプレート）を用いて、60個の無作為に選択された候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図9B】ミラー非小結節テンプレート（108枚の小結節および356枚の非小結節テンプレート）を用いて、60個の無作為に選択された候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図9C】小結節および非小結節の両方にミラーテンプレート（216枚の小結節および356枚の非小結節テンプレート）を用いて、60個の無作為に選択された候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図10】小結節テンプレートのスケーリングを用いて、60個の無作為に選択された候補について、432枚の小結節テンプレートおよび178枚の非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図11A】小結節テンプレートの回転を用いて、60個の無作為に選択された候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図11B】非小結節テンプレートの回転を用いて、60個の無作為に選択された候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図11C】小結節テンプレートおよび非小結節テンプレートの回転を用いて、60個の無作為に選択された候補について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図12A】テンプレートのマトリックスサイズが $24 \times 24$ 画素に等しいときに、60個の無作為に選択された候補について、216枚の小結節テンプレートおよび356枚の非小結節テンプレートの最大相互相関値間の関係を示すグラフ。

【図12B】テンプレートのマトリックスサイズが $36 \times 36$ 画素に等しいときに、60個の無作為に選択された候補について、216枚の小結節テンプレートおよび356枚の非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図12C】テンプレートのマトリックスサイズが $48 \times 48$ 画素に等しいときに、60個の無作為に選択された候補について、216枚の小結節テンプレートおよび356枚の非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図13】小結節状非小結節テンプレートの例を示す図。

【図14】小結節状非小結節テンプレートを除去する前(X)および後(丸)で、試験セットの44個の小結節について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

10

【図15】小結節状非小結節テンプレートを除去する前(X)および後(点)で、試験セットの377個の非小結節の半分(189)について、小結節テンプレートおよび非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

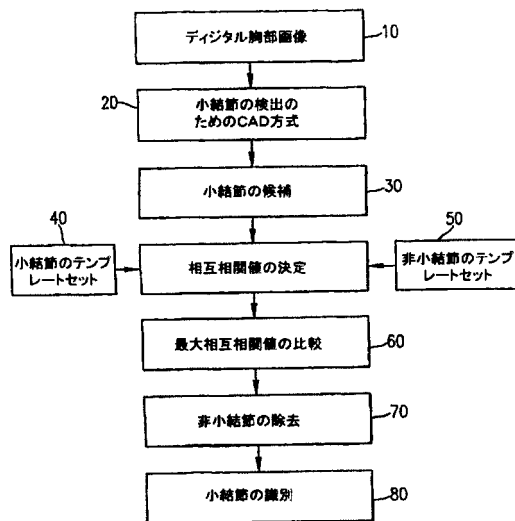
【図16】検証テストで小結節状非小結節テンプレートを除去した後で、試験セットの44個の小結節(円)および377個の非小結節(点)について、5,664枚の小結節テンプレートおよび18,462枚の非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

【図17】66個のCT小結節(丸)および60個のCT非小結節(点)について、1,440枚のCT小結節テンプレートおよび1,440枚のCT非小結節テンプレートとの最大相互相関値間の関係を示すグラフ。

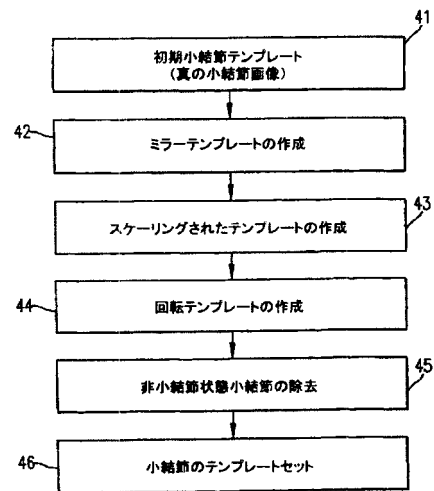
20

【図18】10個の悪性小結節(丸)および10個の良性小結節(点)について、180枚の悪性小結節テンプレートおよび180枚の良性小結節テンプレートとの最大相互相関値間の関係を示すグラフ。180枚の悪性小結節テンプレートおよび180枚の良性小結節テンプレートは、ミラーテンプレート、スケーリングしたテンプレート、および回転したテンプレートを用いて、10個の典型的な悪性小結節および10個の典型的良性小結節からそれぞれ得たものであり、破線は悪性小結節と良性小結節を見分けるための閾値として使用した。

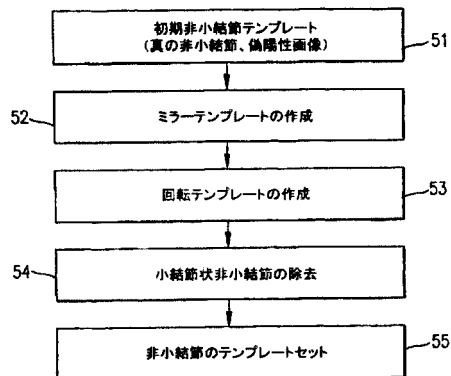
【図1】



【図2】



【図3】



【図5】



【図6】



【図4】

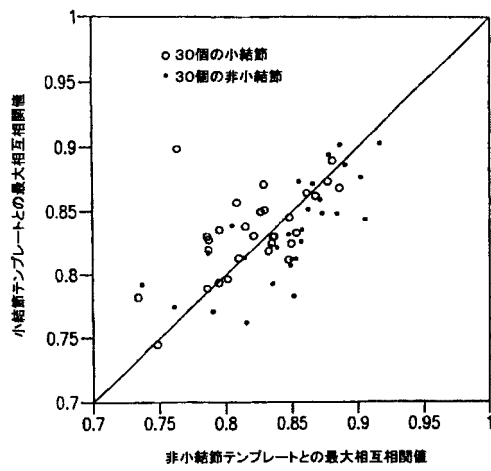


【図7】

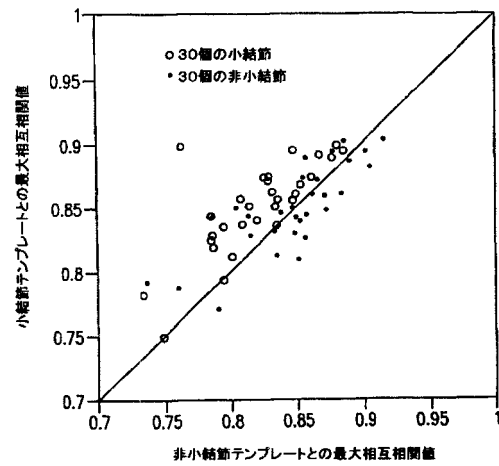




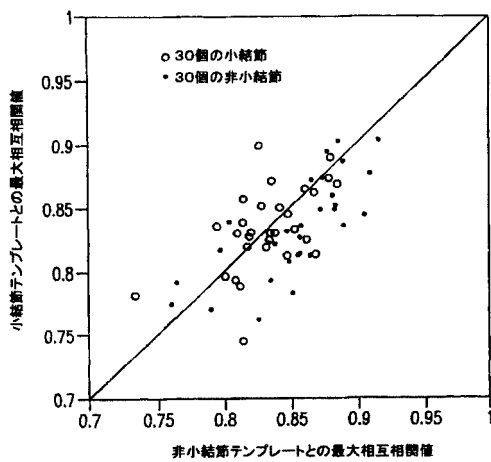
【図 8】



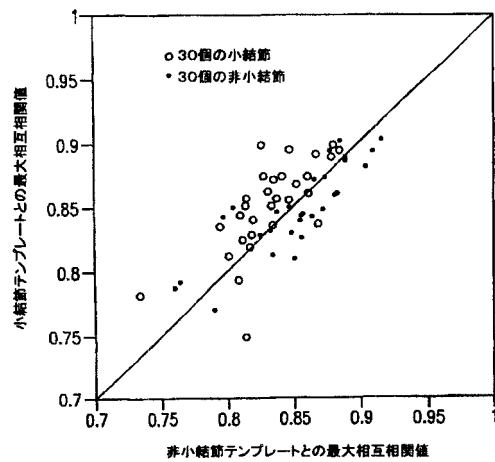
【図 9 A】



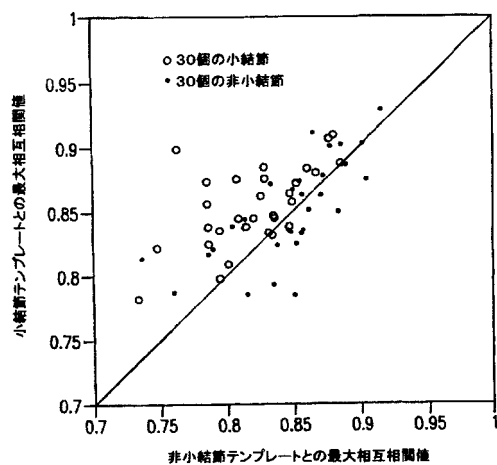
【図 9 B】



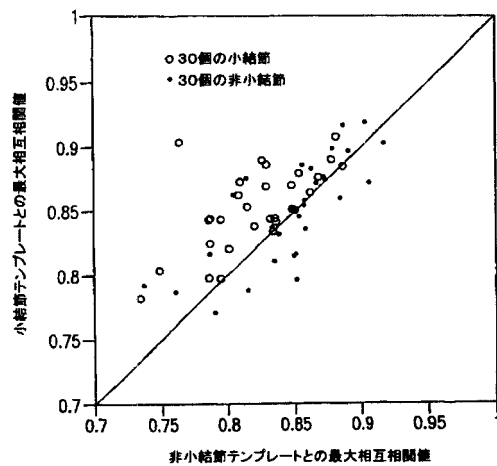
【図 9 C】



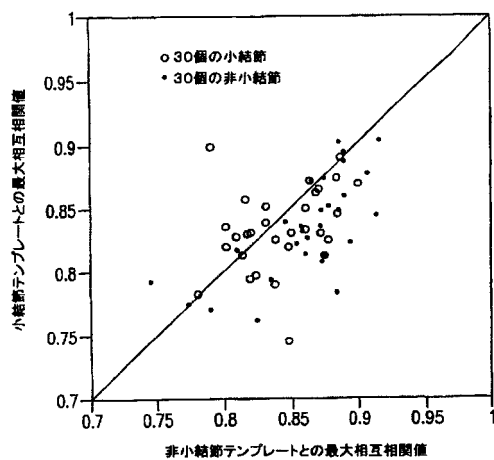
【図 10】



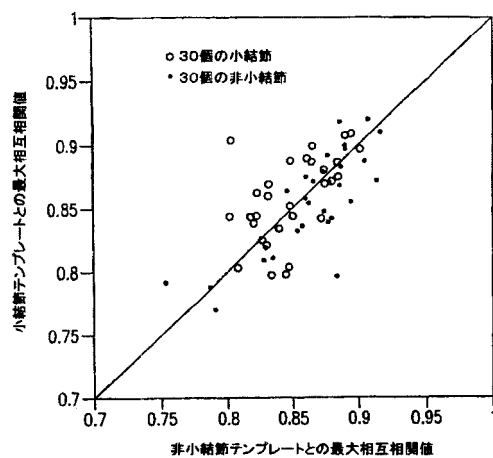
【図 11 A】



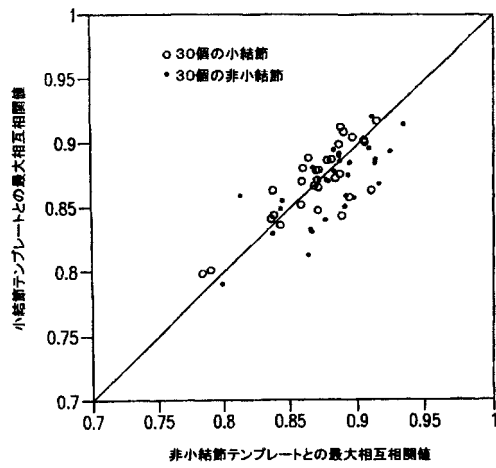
【図 11 B】



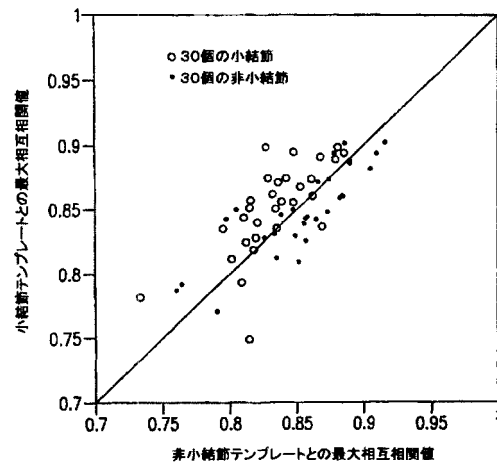
【図 11 C】



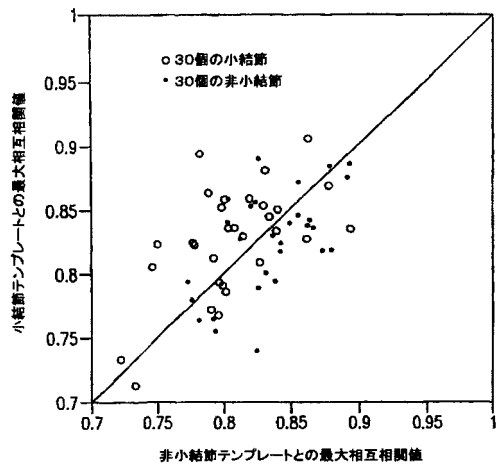
【図 1 2 A】



【図 1 2 B】



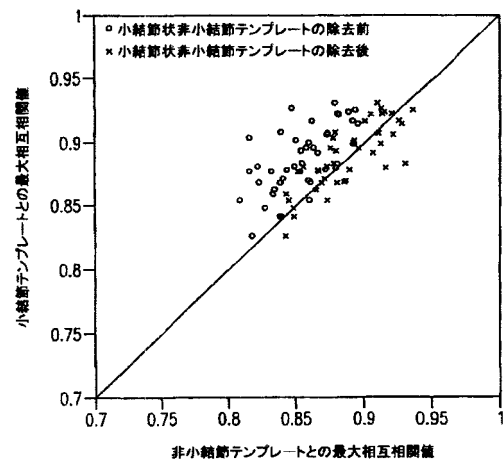
【図 1 2 C】



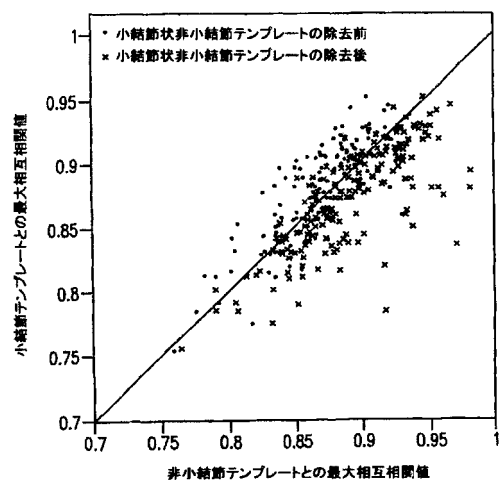
【図 1 3】



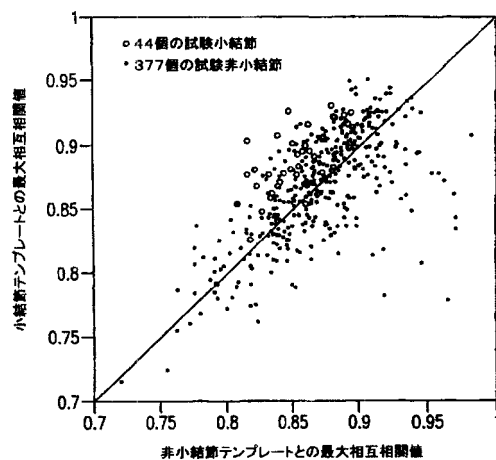
【図 1 4】



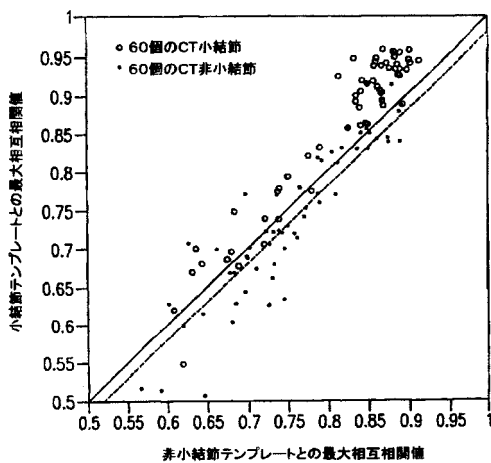
【図 15】



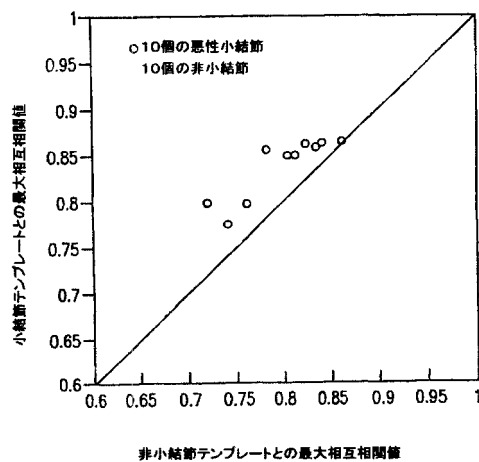
【図 16】



【図 17】



【図 18】



## 【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
6 June 2002 (06.06.2002)

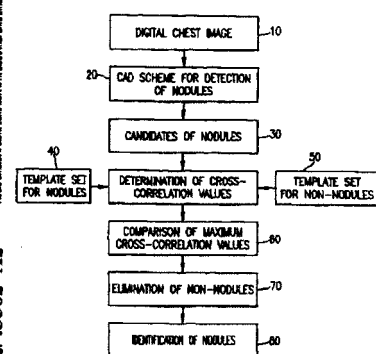
PCT

(10) International Publication Number  
WO 02/43562 A2

- (51) International Patent Classification: A61B 21 8 Nozomigawa, Ots 870-0270 (JP), DOI, Kusto, 6415 Lane Court, Willowbrook, IL 60521 (US).
- (21) International Application Number: PCT/US01/43146
- (22) International Filing Date: 21 November 2001 (21.11.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/716,335 21 November 2000 (21.11.2000) US
- (71) Applicant: ARCH DEVELOPMENT CORPORATION (US/US); 5640 South Ellis Avenue, Room 405, Chicago, IL 60637 (US).
- (72) Inventors: Li, Qiang, 5730 Concord Lane #15, Clarendon Hills, IL 60514 (US); KATSURAGAWA, Shigehiko;
- (74) Agent: KUESTERS, Eckhard, H., Ohlen, Spivak, McClelland, Maier & Neustadt, P.C., 1755 Jefferson Davis Highway, Fourth Floor, Arlington, VA 22202 (US).
- (81) Designated States (national): AB, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CY, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, DE, DK, ES, FI, FR,

[Continued on next page]

(54) Title: PROCESS, SYSTEM AND COMPUTER READABLE MEDIUM FOR PULMONARY NODULE DETECTION USING MULTIPLE TEMPLATES MATCHING

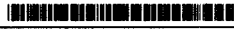


(57) Abstract: A method to determine whether a candidate abnormality in a medical digital image is an actual abnormality, a system which implements the method, and a computer readable medium which stores program steps to implement the method, wherein the method includes obtaining a medical digital image including a candidate abnormality, obtaining plural first templates and plural second templates respectively corresponding to predetermined abnormalities and predetermined non-abnormalities; comparing the candidate abnormality with the obtained first and second templates to derive cross-correlation values between the candidate abnormality and each of the obtained first and second templates; determining the largest cross-correlation value derived in the comparing step and whether the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates or with the second templates, and determining the candidate abnormality to be an actual abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates.

abnormality with the first templates and determining the candidate abnormality to be a non-abnormality when the largest cross correlation value is produced by comparing the candidate abnormality with the second templates. An actual abnormality is similarly classified as malignant or benign based on further cross-correlation values obtained by comparisons with additional templates corresponding to malignant and benign abnormalities.

WO 02/43562 A2

WO 02/43562 A2



GB, GR, IB, IT, LU, MC, NL, PT, SE, TR), OAPI patent  
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,  
NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette*

**Published:**

— without international search report and to be republished upon receipt of that report

WO 02/43562

PCT/US01/43146

**Process, System And Computer Readable Medium For Pulmonary Nodule Detection  
Using Multiple-templates Matching**

**Statement Regarding Federally Sponsored Research**

[0001] The present invention was made in part with U.S. Government support under USPHS grant numbers CA62625 and CA64370 (National Institute of Health). The U.S. Government has certain rights in the invention.

**BACKGROUND OF THE INVENTION**

**Field of the Invention**

[0002] This invention relates to a process, system and computer readable medium for the automated detection of pulmonary nodules in medical images.

[0003] The present invention also generally relates to computerized techniques for automated analysis of digital images, for example, as disclosed in one or more of U.S. Patents 4,839,807; 4,841,555; 4,851,984; 4,875,165; 4,907,156; 4,918,534; 5,072,384; 5,133,020; 5,150,292; 5,224,177; 5,289,374; 5,319,549; 5,343,390; 5,359,513; 5,452,367; 5,463,548; 5,491,627; 5,537,485; 5,598,481; 5,622,171; 5,638,458; 5,657,362; 5,666,434; 5,673,332; 5,668,888; 5,740,268; 5,790,690; 5,832,103; 5,873,824; 5,881,124; 5,931,780; 5,974,165 (PCT Publication WO 95/15537); 5,982,915; 5,984,870; 5,987,345; 6,011,862; 6,058,322; 6,067,373; 6,075,878; 6,078,680; 6,088,473; 6,112,112; 6,138,045; and 6,141,437, as well as U.S. patent applications 08/173,935; 08/900,188; 08/900,189; 08/979,639; 08/982,282; 09/027,468; 09/028,518; 09/092,004; 09/121,719; 09/141,535; 09/298,852 and 09/471,088; and U.S. provisional patent applications 60/107,095; 60/160,790; 60/176,297; 60/176,304; 60/180,162; 60/193,072 and 60/207,401, all of which are incorporated herein by reference.

[0004] The present invention includes use of various technologies referenced and described in the above-noted U.S. patents and patent applications, as well as described in the references identified in the appended APPENDIX by the author(s) and year of publication and cross-referenced throughout the specification by bold numerals in brackets corresponding to the respective references listed in the APPENDIX, the entire contents of which, including the related patents and applications listed above and references listed in the APPENDIX, are incorporated herein by reference.

WO 02/43562

PCT/US01/43146

Discussion of the Background

[0005] It has been reported that radiologists can fail to detect pulmonary nodules on chest radiographs in as many as 30% of positive cases. [1, 2] Many of the lung cancers missed by radiologists were actually visible in retrospect on previous radiographs. [3] Therefore, the inventors and others at the University of Chicago Department of Radiology have developed a computer-aided diagnostic (CAD) scheme to assist radiologists in the detection of pulmonary nodules on digital chest radiographs. [4-9] One problem with the pre-existing scheme is the relatively large number of false positives produced by the automated scheme, which constitutes a main difficulty in the clinical application of the CAD scheme for nodule detection.

SUMMARY OF THE INVENTION

[0006] Accordingly, the object of this invention is to provide CAD process, system and computer program product whereby the number of false positives that are incorrectly reported as nodules is reduced.

[0007] This and other objects are achieved according to the present invention by providing a new and improved method to determine whether a candidate abnormality in a medical digital image is an actual abnormality, a system which implements the method, and a computer readable medium which stores program steps to implement the method, wherein the method includes obtaining plural first templates and plural second templates respectively corresponding to predetermined abnormalities and predetermined non-abnormalities; comparing the candidate abnormality with the obtained first and second templates to derive cross-correlation values between the candidate abnormality and each of the obtained first and second templates; determining the largest cross-correlation value derived in the comparing step and whether the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates or with the second templates; and determining the candidate abnormality to be an actual abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates and determining the candidate abnormality to be a non-abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the second templates. An actual abnormality is similarly determined to be malignant or benign based on further cross-



WO 02/43562

PCT/US01/43146

correlation values obtained by comparisons with additional templates corresponding to malignant and benign abnormalities.

[0008] The maximum cross-correlation values obtained with nodule templates and with non-nodule templates for each of the candidate nodules are employed for distinguishing non-nodules from nodules because a nodule is generally more similar to nodule templates than to non-nodule templates, and a non-nodule is more similar to non-nodule templates than to nodule templates. Therefore, the maximum cross-correlation value of a nodule with nodule templates is generally greater than that with non-nodule templates, and vice versa.

Accordingly, according to the present invention, the greatest cross-correlation value obtained is determined and the candidate nodule is then determined to be an actual nodule when the greatest cross-correlation value is obtained based on a comparison with a nodule template and to be a false positive when the greatest correlation value is obtained based on a comparison with a non-nodule template.

[0009] A study implementing the CAD process of the invention was performed, whereby a large number of false positives (44.3%) in chest radiographs were removed with reduction of a very small number of true positives (2.3%) by use of the multiple-templates matching technique. In addition, a similar result on another CAD scheme for detection of nodules on CT images by use of the multiple-templates matching technique was obtained. Thus, the present invention is considered to have applicability to improve the performance of many different CAD schemes for detection of various lesions in medical images, including nodules in chest radiographs, masses and microcalcifications in mammograms, nodules, colon polyps, liver tumors, and aneurysms in CT images as well as breast lesions in ultrasound and magnetic resonance images. Furthermore, the multiple-templates matching technique has application to distinguish benign lesions from malignant lesions, in order to improve the performance of CAD schemes for classification between benign lesions and malignant lesions such as lung cancer, breast cancer, colon cancer and stomach cancer.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0010] A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by

WO 02/43562

PCT/US01/43146

reference to the following detailed description when considered in connection with the accompanying drawings wherein:

[0011] Figure 1 is a flow chart illustrating an overall multiple-templates matching process according to the present invention.

[0012] Figure 2 is a flow chart illustrating a basic process for the creation of a nodule template set.

[0013] Figure 3 is a flow chart illustrating a basic process for the creation of a non-nodule template set.

[0014] Figure 4 is an illustration of a plurality of examples of nodule templates which are corrected for background trend.

[0015] Figure 5 is an illustration of a plurality of examples of non-nodule templates which are corrected for background trend.

[0016] Figure 6 is an illustration of a plurality of examples of original images for nodule templates.

[0017] Figure 7 is an illustration of a plurality of examples of estimated background images obtained by fitting a linear function to the original images in Fig. 6.

[0018] Figure 8 is a graph illustrating the relationship between the maximum cross-correlation values with 108 nodule templates and 178 non-nodule templates, for 60 randomly selected candidates.

[0019] Figures 9 (a), 9(b) and 9(c) are graphs illustrating the relationship between the maximum cross-correlation values with nodule templates and non-nodule templates, for 60 randomly selected candidates with (a) mirror nodule templates (216 nodule and 178 non-nodule templates), (b) mirror non-nodule templates (108 nodule and 356 non-nodule template), and (c) mirror templates for both nodules and non-nodules (216 nodule and 356 non-nodule template), respectively

[0020] Figure 10 is a graph illustrating the relationship between the maximum cross-correlation values with 432 nodule templates and 178 non-nodule templates, for 60 randomly selected candidates, with scaling of the nodule templates.

[0021] Figures 11(a), 11(b) and 11(c) are graphs illustrating the relationship between the maximum cross-correlation values with nodule templates and non-nodule templates, for 60

WO 02/43562

PCT/US01/43146

randomly selected candidates with (a) rotation of nodule templates (324 nodule and 178 non-nodule templates), (b) rotation of non-nodule templates (108 nodule and 534 non-nodule templates), and (c) rotation of the nodule and non-nodule templates (324 nodule and 534 non-nodule templates), respectively.

[0022] Figures 12(a), 12(b) and 12(c) are graphs illustrating the relationship between the maximum cross-correlation values with 216 nodule templates and 356 non-nodule templates, for 60 randomly selected candidates when the matrix size of templates is equal to (a) 24x24, (b) 36x36, and (c) 48x48 pixels, respectively.

[0023] Figure 13 is an illustration of examples of nodule-like non-nodule templates.

[0024] Figure 14 is a graph illustrating the relationship between the maximum cross-correlation values with nodule templates and non-nodule templates, for the 44 nodules in the test set, before (x's) and after (circles) the elimination of nodule-like non-nodule templates.

[0025] Figure 15 is a graph illustrating the relationship between the maximum cross-correlation values with nodule templates and non-nodule templates, for half (189) of the 377 non-nodules in the test set, before (x's) and after (dots) the elimination of nodule-like non-nodule templates.

[0026] Figure 16 is a graph illustrating the relationship between the maximum cross-correlation values with 5,664 nodule templates and 18,462 non-nodule templates, for 44 nodules (circles) and 377 non-nodules (dots) in the test set, after the elimination of nodule-like non-nodule templates in a validation test.

[0027] Figure 17 is a graph illustrating the relationship between the maximum cross-correlation values with 1,440 CT nodule templates and 1,440 CT non-nodule templates, for 60 CT nodules (circles) and 60 CT non-nodules (dots).

[0028] Figure 18 is a graph illustrating a relationship between the maximum cross-correlation values with 180 malignant nodule templates and 180 benign nodule templates, for 10 malignant nodules (circles) and 10 benign nodules (dots), wherein the 180 malignant nodule templates and 180 benign nodule templates were obtained, respectively, from the 10 typical malignant nodules and 10 typical benign nodules by use of mirror templates, scaled templates, and rotated templates and the dashed line was used as a threshold for distinction between malignant nodules and benign nodules.

WO 02/43562

PCT/US01/43146

DESCRIPTION OF THE PREFERRED EMBODIMENTSMethods and Materials

[0029] The chest images used in first development of the present invention, as directed to chest radiographs, consisted of 100 abnormal posteroanterior (PA) chest radiographs selected in the Department of Radiology, the University of Chicago Hospitals [9]. A total of 122 nodules were confirmed, based on the consensus of two radiologists and verified by CT scans or radiographic follow-up. The nodule diameter ranged from 4 mm to 27 mm (average, 13 mm). The digital images were obtained by digitizing of the chest radiographs with a Konica laser film scanner. The original digital image has a pixel size of 0.175 mm, a matrix size of 2000x2000, and a gray level range of 10 bits. Because nodules are relatively large in chest images, the matrix size of the digital images was reduced by a factor of 4. Consequently, the final image had a pixel size of 0.7 mm and a matrix size of 500x500. In addition to the main database above, a supplemental database, the Japanese Standard Digital Image Database [10], was employed solely for creation of nodule templates and for training of the multiple-templates matching technique, but not for the verification of the present CAD scheme with the multiple-templates matching technique. The supplemental database was developed by the Japanese Society of Radiological Technology, and is described hereinafter.

[0030] With the pre-existing CAD scheme, 116 nodules and 4875 false positives were initially detected from the 100 chest images in our database. A rule-based classification technique was then applied to reduce the number of false positives, [9] with which 88 nodules and 377 false positives were obtained. For the multiple-templates matching technique, the 116 nodules and 4875 false positives detected initially by the pre-existing CAD scheme were used as initial nodule templates and initial non-nodule templates, respectively. However, the 88 nodules and 377 false positives reported finally by the present CAD scheme were used for training and testing of the final result. Eight non-nodule-like nodule templates and 1798 nodule-like non-nodule templates were systematically eliminated from the nodule template set and the non-nodule template set, respectively, so as to achieve a good performance for the multiple-templates matching technique. It is also noted that the 377 false positives in the test candidate set survived the pre-existing rule-based tests and are, therefore, considered as "difficult" false positives, which are similar to nodules. In the study, the multiple-templates

WO 02/43562

PCT/US01/43146

matching technique was employed to reduce significantly the number of these difficult false positives.

Summary of the pre-existing CAD scheme based on rule-based tests

[0031] First of all, the lung areas in PA chest images were segmented by use of the delineated ribcage edges, lung top, and diaphragm, [11, 12] and were employed for the subsequent processing of the CAD scheme. A difference image was then obtained by subtraction of a nodule-suppressed image (by use of a smooth filter) from a nodule-enhanced image (by use of a matched filter), so that the complicated background structure could be reduced, and thus nodules could be more conspicuous. In order to detect the initial nodule candidates from the difference image, multiple binary images were obtained by thresholding the difference image with various threshold levels. In each of the binary images, a component labeling technique was used to identify each isolated "island," [4] and two features, the effective diameter and the degree of circularity, were determined for each island. An island was then considered to be an initial nodule candidate if its effective diameter and degree of circularity were equal to or larger than 6.5 mm and 0.65, respectively. For the database of 100 abnormal chest images, a total of 116 (out of 122) nodules and 4875 false positives were identified as initial nodule candidates.

[0032] Next, a region growing technique [5, 6] was applied to both the difference image and the original image at the locations of initial nodule candidates, for accurate segmentation of nodule candidates from background in each of the two images. Various features, such as the effective diameter, contrast, degree of circularity, degree of irregularity, edge gradient, slope of circularity, slope of irregularity, and slope of diameter, were then determined from each of the grown regions, and employed for distinction between nodules and false positives by a rule-based classification techniques. After this step, most nodules, 88 (75.9%) out of 116, were retained, and most false positives, 4498 (92.3%) out of 4875, were eliminated; thus, 377 false positives remained.

Overall scheme of the multiple-templates matching technique for reduction of false positives

WO 02/43562

PCT/US01/43146

[0033] Figure 1 shows the overall scheme for the multiple-templates matching process of the present invention. First of all, in step 10 a digital chest image is obtained and in step 20 a CAD scheme for nodule detection is applied and in step 30 candidate nodules are obtained thereby. In steps 40 and 50, nodule template sets and non-nodule template sets are obtained, and in step 60, for each of the candidate nodules, a cross-correlation technique is employed to calculate the cross-correlation values with the nodule template set and with the non-nodule template set. Finally, two maximum cross-correlation values obtained with the nodule template set and with the non-nodule template set are determined, and employed to eliminate false positives (non-nodules) in step 70. Resulting in the identification of the remaining candidate nodules as actual nodules in step 80. If the maximum cross-correlation value is obtained by comparison of the candidate nodule with a non-nodule template, then the candidate nodule is determined to be a false positive. On the other hand, if the maximum cross-correlation value is obtained by comparison of the candidate nodule with a nodule template, then the candidate nodule is determined to be an actual nodule.

[0034] In order to apply the multiple-templates matching technique, two multiple-template sets, i.e., one with a large number of nodule templates and another with a large number of non-nodule templates, were created. In this study, the initial nodule candidates (116 nodules and 4875 non-nodules) reported by the pre-existing CAD scheme were used as initial templates, which were small regions of interest (ROIs) of 36x36 pixels centered at the locations of the initial nodule candidates.

[0035] Figure 2 shows the basic scheme for creation of the nodule template set from the initial nodule templates. In step 41 the initial nodule templates were obtained and then in step 42, each of the nodule templates was first right/left reversed to produce a mirror template so that the number of templates would be doubled. Then in step 43, each of the nodule templates was scaled (minified or magnified), and in step 44 also rotated to increase the number of nodule templates further. Finally, in step 45 eight non-nodule-like nodule templates (atypical nodules such as very subtle nodules overlapping ribs) were carefully identified as inappropriate nodule templates, and thus excluded from the nodule template set, to derive in step 46 the template set for nodules.

WO 02/43562

PCT/US01/43146

[0036] Figure 3 shows a similar scheme for creation of the non-nodule template set. Each of the non-nodule templates obtained in step 51 was first right/left reversed in step 52, and also rotated in step 53 to increase the number of non-nodule templates. Many (1789) non-nodule templates that had a similar appearance as nodules were then systematically removed in step 54 from the non-nodule template set by use of a training set, as will be hereinafter, because the presence of these nodule-like non-nodule templates can considerably degrade the performance of the multiple-templates matching technique. The non-nodule templates were not scaled because a relatively large number of non-nodule templates had been obtained, and also because the templates obtained by use of scaling were slightly obscured. After the removal of non-nodule-like nodule templates and nodule-like non-nodule templates, 3077 non-nodule templates in the non-nodule template set were obtained in step 55, in addition to the 108 nodule templates obtained in the nodule template set.

[0037] Figures 4 and 5 show typical examples of nodule templates and non-nodule templates, respectively. It is apparent in Figs. 4 and 5 that nodule templates and non-nodule templates are quite different in appearance, and are thus the basis for distinction between nodules and non-nodules.

[0038] The 88 nodules reported finally by the pre-existing CAD scheme were then divided into two sets, each of which had 44 randomly selected nodules. One set was employed, which is called a training set, to systematically eliminate nodule-like non-nodule templates from the non-nodule template set, as described hereinafter. The 44 nodules in the other set plus the 377 non-nodules (false positives), called a test set, were used for verifying the performance of the multiple-templates matching technique of the present invention. Similar to the templates, a test candidate was a small ROI of 40x40 pixels centered at the location of nodule candidates in the test set.

[0039] In addition to the data sets above, a small template set and a small test set were also used, in order to reduce the computation time for determination of the matrix size of templates, and for investigation of the effect of a number of parameters such as scaling and rotation of templates. The small template set included all of the 108 nodule templates in the nodule template set and 178 "typical" non-nodule templates selected manually from 4875

WO 02/43562

PCT/US01/43146

non-nodules. The small test set included 30 nodules and 30 non-nodules, which were randomly selected from the 88 nodules and 377 non-nodules, respectively.

Determination of maximum cross-correlation values between a test candidate and the template sets of nodules and non-nodules

[0040] A round robin method, i.e., a leave-one-out test method, was used for the determination of cross-correlation values between templates and a test candidate; namely, if a test candidate was included in the set of templates, the corresponding template would not be used for the determination of the cross-correlation value with the test candidate. Before the actual calculation of the cross-correlation value between a template and a test candidate, a preprocessing step was utilized for correction of the background trend included in a template and test candidate. The background trend in the template (or test candidate) was represented by a two-dimensional (surface) linear function, and the coefficients of the linear function were determined by a least square method. The estimated surface function was then subtracted from the original image of the template (or the test candidate) to provide a background-trend corrected image.

[0041] Figures 6 and 7 show the original images for templates and the estimated background, respectively. The background-trend corrected images, which were obtained by subtraction of the estimated background from the original images, are shown in Fig. 4.

[0042] The matrix size (40x40 pixels) of the test candidates is 4 pixels larger than that (36x36 pixels) of the templates in both horizontal and vertical directions. In order to determine the cross-correlation value between a template  $A$  and a test candidate  $B$ , template  $A$  was moved for a maximum shift value of 4 pixels in both horizontal and vertical directions, and a cross-correlation value  $C_{ij}$  was calculated at each shift value ( $i,j$ ) by the equation:

$$C_{ij}^2 = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \frac{\{A(m,n) - \bar{A}\} \{B_{ij}(m,n) - \bar{B}\}}{\sigma_A \sigma_B}, \quad i,j \in \{0,1,2,3,4\},$$

where  $\bar{A}$  and  $\sigma_A^2$  are the mean and variance of the pixel values in the template  $A$ , respectively, and  $\bar{B}$  and  $\sigma_B^2$  are the mean and variance of the pixel values in a subregion,  $B_{ij}$ .



WO 02/43562

PCT/US01/43146

of the test candidate B at a shift value  $(i,j)$ , respectively. The mean and variance of the pixel values in regions A and  $B_{ij}$  are defined by the following equations:

$$\bar{A} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N A(m,n),$$

$$\sigma_A^2 = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \{A(m,n) - \bar{A}\}^2,$$

$$\bar{B} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N B_{ij}(m,n),$$

The largest cross-correlation value among all shift

$$\sigma_A^2 = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \{A(m,n) - \bar{A}\}^2,$$

values was then determined as the cross-correlation value between the template and the test candidate. Next, the cross-correlation values for the test candidate with all of the nodule templates were computed, and the maximum cross-correlation value was determined, and employed as a unique feature of the test candidate, which indicates the extent of the resemblance to the nodule. Similarly, the maximum cross-correlation value for the test candidate with the non-nodule template set was determined and used as another unique feature of the test candidate, which indicates the extent of the resemblance to the non-nodule. These two features for the candidates in the test set were then employed for distinguishing false positives from nodules, when the maximum cross-correlation values with the non-nodule templates was larger than that with the nodule templates.

#### Creation of nodule templates and non-nodule templates

[0043] A key factor for the success of the multiple-templates matching technique is the number of templates available for nodules and non-nodules. In this study, only 108, which

WO 02/43562

PCT/US01/43146

were selected from the 116 nodules initially detected by our CAD scheme for nodule detection nodules, were used as nodule templates. Although any ROIs in a chest radiograph that do not contain a nodule can theoretically be considered as a non-nodule template, as non-nodule templates, the 3077 false positives were selected from the 4875 false positives initially detected by the preexisting CAD scheme because they were considered to be "typical" non-nodules. The following three methods were then utilized to increase the number of templates in this study:

- (1) Right/left reversing of a template to create a mirror template,
- (2) scaling of a template by three different factors of 0.6, 0.8, and 1.2,
- (3) rotation of a template by two different angles, -10 and +10 degrees.

The ROIs obtained directly from the original images were called original templates, and those obtained by right/left reversing, scaling, or rotation of an original template were called derived templates. It should be noted that additional derived templates can be created by scaling with additional factors and also by rotation with additional angles.

[0044] The right/left reversing enables a template in the right lung to be useful in the left lung and vice versa, and it doubles the number of templates. The scaling and the rotation in this study increased the number of templates by factors of four and three, respectively, and thus the total number of templates can be increased 24 times by a combination of all of these methods.

[0045] Figure 8 shows the relationship between the maximum cross-correlation values with 108 nodule templates and 178 non-nodule templates in the small template set for 60 candidates in the small test set. As has been described previously, the templates were carefully selected from typical nodules and non-nodules. It is apparent in Fig. 8 that, although there is considerable overlap, nodules tend to have larger maximum correlation values with nodule templates, whereas non-nodules do so with non-nodule templates, i.e., nodules tend to be above the 45-degree line and non-nodules below the 45-degree line. This general trend indicates the usefulness of the multiple-templates matching technique in distinguishing between nodules and non-nodules. However, many test candidates have relatively low cross-correlation values with either nodule or non-nodule templates, which

WO 02/43562

PCT/US01/43146

indicates that there were not enough templates, in the template set for nodules or non-nodules, which are very similar to those test candidates.

[0046] Figure 9 shows the relationship between the maximum cross-correlation values with the nodule templates and non-nodule templates for the 60 test candidates, when additional templates were incorporated based on (a) mirror nodule templates only, (b) mirror non-nodule templates only, and (c) mirror templates for both nodules and non-nodules. When Fig. 8 and Fig. 9(a) are compared, it is clear that most test candidates have moved upward in Fig. 9(a), and thus the maximum cross-correlation values with nodule templates for these test candidates were increased by the use of mirror templates for nodules. In addition, nodules are located on or above the diagonal line in Fig. 9(a), and thus some non-nodules below the diagonal line can be clearly distinguished from nodules. Similarly, it was observed that most test candidates have moved to the right in Fig. 9(b), compared with Fig. 8, and thus the use of mirror templates for non-nodules increased the maximum cross-correlation values with non-nodule templates. It is apparent in Fig. 9(c) that the test candidates have moved upward and to the right, and the maximum cross-correlation values with both nodule and non-nodule templates were increased by use of mirror templates.

[0047] Although the scaling may obscure a template in some cases, it was still applied to the nodule templates because the study had only a very limited number of nodules. Figure 10 demonstrates the relationship between the maximum cross-correlation values with the nodule templates and non-nodule templates, together with scaled nodule templates. When the results in Fig. 10 are compared with the results without scaled nodule templates in Fig. 8, it is apparent that the addition of scaled nodule templates resulted in improved maximum cross-correlation values with nodule templates for the test candidates, and also improved distinction between nodules and non-nodules, because most nodules are above the diagonal line.

[0048] On the other hand, unlike scaling, the rotation was applied to all of the nodule and non-nodule templates. Figure 11 shows the relationship between the maximum cross-correlation values with the nodule templates and non-nodule templates, when additional templates were created by (a) rotation of the nodule templates only, (b) rotation of the non-nodule templates only, and (c) rotation of both nodule and non-nodule templates; results without rotated templates are shown in Fig. 8. Again, the maximum cross-correlation values

WO 02/43562

PCT/US01/43146

with the nodule templates and non-nodule templates became larger by addition of rotated templates.

[0049] Another parameter for the multiple-templates matching technique is an appropriate choice of the matrix sizes for the templates and the test candidates. In this study, the effect of various matrix sizes ranging from 24x24 pixels to 48x48 pixels on the overall performance was examined. It was found that the best results were obtained when the matrix sizes for the templates and the test candidates were 36x36 pixels (approximately 25x25 mm<sup>2</sup>) and 40x40 pixels (28x28 mm<sup>2</sup>), respectively.

[0050] Figures 12(a), (b), and (c) show the relationship between the maximum cross-correlation values with the nodule templates and non-nodule templates for 60 candidates, when the matrix size of templates was 24x24, 36x36, and 48x48 pixels, respectively. The templates used included 108 nodules and 178 non-nodules, together with their mirror templates. Although the maximum cross-correlation values were generally large, with a matrix size of 24x24 pixels, the overlap of the correlation values between the nodules and non-nodules was large as well, which is not useful for the separation of nodules and non-nodules. On the other hand, when the matrix size was 48x48 pixels, the maximum cross-correlation values were usually low and could not be used as reliable features. It is apparent from Fig. 12 that the best separation between the 30 nodules and 30 non-nodules was achieved when the matrix size of the templates was 36x36 pixels. Therefore, the matrix size of the templates used in this study was determined as 36x36 pixels.

#### Elimination of nodule-like non-nodules in non-nodule template set

[0051] It is important to note that not all non-nodule templates can make a useful contribution to the improvement in the performance of the multiple-templates matching technique. In fact, many non-nodule templates do impair the performance of the multiple-templates matching technique. As can be seen in Fig. 9(c), there are 4 nodules located below the diagonal line, which implies that these nodules are more similar to some non-nodule templates used than to all of the nodule templates. Figure 13 illustrates examples for such non-nodule templates that pulled some nodules below the diagonal line by increasing the maximum correlation values of nodules with non-nodule templates. It was found that these

WO 02/43562

PCT/US01/43146

non-nodule templates typically resembled nodules in appearance. Therefore, it is desirable to eliminate these nodule-like non-nodule templates in order to achieve a good performance for the multiple-templates matching technique. The nodules in the training set and in a supplemental database were employed to achieve this task.

[0052] In addition to the main database described previously, another supplemental database was also employed solely for creation of more nodule templates and for training of the multiple-templates matching technique, but not for the verification of the present CAD scheme with the multiple-templates matching technique, because the characteristics of the chest images in the supplemental database are quite different from those in the main database. The supplemental database included 128 chest images with solitary lung nodules, which were selected from a total of 154 nodule cases in the Japanese Standard Digital Image Database developed by the Japanese Society of Radiological Technology. [10] Twenty-six nodule cases were eliminated from the Japanese standard database, each of which contained a nodule with a subtlety rating score of either one or five, corresponding to an extremely subtle nodule or an obvious nodule, respectively. The original chest images were digitized with a 0.175 mm pixel size, a matrix size of 2048x2048, and 12-bit gray levels. In this study, the matrix size was reduced to 512x512 by subsampling of the original image data by a factor of 4, and the number of gray levels was decreased to 10 bits, in order to be consistent with the chest images in the main database. The 128 nodules in the supplemental database together with the 108 nodules in the main database were used as the nodule template set hereafter for verification of the multiple-templates matching technique. The 128 nodules were also employed for training of the multiple-templates matching technique, namely, for the removal of nodule-like non-nodule templates.

[0053] As the first step in training for the multiple-templates matching technique, for each of the 128 nodules, twenty non-nodule templates which provided the 20 largest cross-correlation values with the nodule were considered here to be nodule-like non-nodules, and were eliminated from the 4,875 original templates in the initial non-nodule template set. A total of 1,338 original non-nodule templates were thus removed, and 3,537 original non-nodule templates remained in the non-nodule template set. Similarly, the 44 nodules in the training set were employed for further elimination of nodule-like non-nodules from the 3,537

WO 02/43562

PCT/US01/43146

original templates. Thus, 460 original non-nodule templates were again eliminated, and finally, 3,077 original templates were left in the non-nodule template set. A total of 1,798 templates were removed from the initial non-nodule template set by employing the 128 nodules in the supplemental database and the 44 nodules in the training set altogether. For the removal of nodule-like non-nodules in the template set, the 20 largest cross-correlation values were empirically employed as the threshold. However, it is possible to employ a different number of largest cross-correlation values, depending on the number and characteristics of templates available.

[0054] In order to demonstrate how we improved the performance of the multiple-templates matching technique by elimination of the nodule-like non-nodules, Fig. 14 shows the relationship between the maximum cross-correlation values with the nodule templates and non-nodule templates for the 44 nodules in the test set, before (indicated by x's) and after (indicated by circles) the removal of the nodule-like non-nodule templates. The nodule template set used in Fig. 14 was composed of 5,664 (24x236) templates, which included the 108 nodules in the main database and the 128 nodules in the supplemental database, and their mirror templates, scaled templates, and rotated templates. Note that the non-nodule template set contained 29,250 (6x4,875) templates and 18,462 (6x3,077) templates, respectively, before and after the removal of the nodule-like non-nodule templates, which included the original non-nodule templates, their mirror templates, and rotated templates. It is apparent in Fig. 14 that most nodules are moved significantly to the left by elimination of the nodule-like non-nodule templates, and that all but one nodule are located above the diagonal line, as can be predicted.

[0055] It should be noted that the elimination of nodule-like non-nodule templates would also affect the non-nodules in the test set. Figure 15 shows the relationship between the maximum cross-correlation values with the nodule templates and non-nodule templates, for one half (189) of the 377 non-nodules in the test set, before (indicated by x's) and after (indicated by dots) the elimination of the non-nodule templates. Note that there were 5,664 templates in the nodule template set, and 29,250 and 18,462 templates in the non-nodule template set, respectively, before and after the elimination of nodule-like non-nodule templates. Only half of the non-nodules are shown in Fig. 15 for a clearer display. The

WO 02/43562

PCT/US01/43146

nodule template set and non-nodule template sets used in Fig. 15 are the same as those employed in Fig. 14. It is apparent in Fig. 15 that, although some of the 377 non-nodules moved to the left as did the nodules in Fig. 14, many non-nodules still remain below the diagonal line after the elimination of the nodule-like non-nodule templates. This findings implies that, after elimination of nodule-like non-nodule templates, it is possible to distinguish the false positives (non-nodules) below the diagonal line in Fig. 15 from the nodules in Fig. 14, because nearly all of the nodules are located above the diagonal lines.

[0056] With the 5,664 nodule templates and 18,462 non-nodule templates after the elimination of nodule-like non-nodules, a validation test was conducted based on the 44 nodules and 377 non-nodules in the test set. It should be noted that the 44 nodules and 377 non-nodules have not been utilized for training of the multiple-templates matching technique, i.e., for elimination of either nodule-like non-nodules or non-nodule-like nodules in the template set, although the 377 non-nodules were shown in Fig. 15 to demonstrate the effect of removing nodule-like non-nodule templates. Figure 16 shows the relationship between the maximum cross-correlation values with the 5,664 nodule templates and 18,462 non-nodule templates for the 44 nodules and 377 non-nodules in the test set. It is apparent in Fig. 16 that a significant distinction can be made between the nodules and false positives (non-nodules) based on the two maximum cross-correlation values. For example, if the diagonal line is used as a threshold, namely, if candidates located above the diagonal line are accepted as nodules, then we can eliminate 167 (44.3%) false positives from the 377 non-nodules with a reduction of only one (2.3%) true nodule. This actually constitutes a significant improvement of the existing CAD scheme because many of the 377 false positives are similar to nodules in appearance, and are thus considered as "difficult" false positives.

[0057] The multiple-templates matching technique requires a relatively large amount of computer time because cross-correlation values are computed with a large number of templates. For example, 24,126 (5,664+18,462) templates were used in the final validation test, which requires about 85 seconds to calculate the cross-correlation values for each of the test candidates on a personal computer with an Intel Pentium-III 733 MHZ CPU and a Linux operating system. In the future, the inventors plan to expand their database greatly in order to achieve a higher performance and reliability, which, in turn, implies a larger template set and

WO 02/43562

PCT/US01/43146

more computer time required for computing the cross-correlation values. However, this will not be a serious problem for the multiple-templates matching technique, because this technique is applied solely in the final step of the CAD scheme, which, on average, reports only three or four nodule candidates per chest image. Moreover, computers are, and will be, becoming faster and faster, and thus will significantly reduce the problem caused by the large computational burden of the multiple-templates matching technique. At present, it takes about 5-6 minutes to process each of the chest images for removal of some false positives by use of the multiple-templates matching technique.

#### Application to Computed Tomography

[0058] In order to demonstrate that the multiple-templates matching technique has the potential to distinguish false positives from nodules in a CAD scheme for nodule detection on computed tomography (CT), a pilot study was conducted to apply the multiple-templates matching technique to a database of 44 CT scans, which contain a total of 237 nodules. The CT scans were obtained with 10 mm collimation, and 10 mm reconstruction intervals. Each CT slice has a matrix size of 512x512 pixels and a gray level range of 10 bits. The field of view was optimized for each patient during the examination so that the pixel sizes in the database ranged from 0.566 to 0.781 mm.

[0059] An existing CAD process [13] at the University of Chicago Department of Radiology for CT nodule detection first segmented lungs from background in each slice by using a thresholding technique and a rolling-ball algorithm. A multiple gray-level thresholding technique was then applied to the segmented lung areas for detection of initial nodule candidates. For each of the initial nodule candidates, six geometric features (volume, sphericity, radius of equivalent sphere, maximum compactness, maximum circularity, and maximum eccentricity) and three gray level features (mean gray level, standard deviation of gray level, and the gray level threshold at which the nodule candidate was first detected) were determined, and were employed for distinguishing false positives from nodules. With this CAD scheme, 208 (87.8%) nodules and 4923 false positives (approximately three false positives per slice) were reported as nodule candidates.



WO 02/43562

PCT/US01/43146

[0060] The 208 nodules and 200 false positives, which were randomly selected from the 4923 false positives, were first chosen for construction of a basis for applying the multiple-templates matching technique to the CT nodule detection scheme. In this study, only those nodules and non-nodules that were distant from the boundaries of lung areas were manually selected, so that the templates obtained from these candidates were completely contained inside the lungs. Sixty nodules and sixty non-nodules were thus obtained from the 208 nodules and 200 non-nodules, respectively, and were used as both templates and test candidates. A leave-one-out (round robin) test method was employed so that a test candidate and the corresponding derived templates were not included in the template set used for testing. With the multiple-templates matching technique described above, the maximum cross-correlation values for 60 nodules (circles) and 60 non-nodules (dots) were obtained by use of 1,440 (24x60) nodule templates and 1,440 (24x60) non-nodule templates, as shown in Fig. 17. The 1,440 nodule templates and 1,440 non-nodule templates were obtained, respectively, from the 60 nodules and 60 non-nodules by use of mirror templates, scaled templates, and rotated templates. The dashed line was used as a threshold for distinction between nodules and non-nodules.

[0061] It is apparent in Fig. 17 that nodules tend to be located above the diagonal line, whereas non-nodules tend to be located below the diagonal line, which indicates that nodules are generally more similar to nodule templates, and non-nodules are generally more similar to non-nodule templates. It is important to note in Fig. 17 that many nodules have very large cross-correlation values with nodule templates, and thus they are distributed above the diagonal line and around the upper-right corner, which implies that each of these nodules has at least one similar nodule template in the template set. It is thus believed that there is a larger probability to find two similar nodules in CT scans than in chest radiographs, because of the simpler background structures in CT scans. If the dashed line is used as a threshold, that is, if those candidates above the dashed line are accepted as nodules and those candidates below the dashed line are accepted as non-nodules, then 29 (48.3%) false positives (non-nodules) can be eliminated with a reduction of only one (1.7%) nodule. This result indicates that the multiple-templates matching technique has the potential to significantly reduce the number of false positives in the CAD scheme for CT nodule detection, and also in many

WO 02/43562

PCT/US01/43146

CAD schemes for detection of many different kinds of lesions, such as masses and microcalcifications in mammograms, breast lesions in ultrasound and magnetic resonance images, colon polyps and liver tumors in abdominal CT images, and aneurysms in brain CT images. In addition, the multiple-templates matching technique can be used to distinguish benign nodules from malignant nodules, in order to improve the performance of CAD schemes for classification between benign lesions and malignant lesions due to many cancers, such as lung cancer, breast cancer, colon cancer, and stomach cancer.

#### Application to Benign Nodules

[0062] A pilot study was also conducted to show that the multiple-templates matching technique can be employed to distinguish benign nodules from malignant nodules in a CAD scheme for nodule classification on digital chest images. A database of 56 chest images, which contained 23 malignant nodules and 33 benign nodules, was used. Each image had a matrix size of 2048x2048 pixels, a pixel size of 0.175 mm, and a gray level range of 10 bits. The location and the size for each of the nodules were identified by three radiologists, and the average values for the location and the size were calculated and used for classification between malignant nodules and benign nodules in chest images, as next discussed.

[0063] 10 "typical" malignant nodules and 10 "typical" benign nodules were manually selected for distinction of benign nodules from malignant nodules by applying the multiple-templates matching technique. Because the approximate nodule size was known, an original template was first obtained at the location of a nodule with a variable matrix size so that the area of the nodule was approximately half of the area of the template. The matrix size for all the templates was then normalized (reduced or magnified) to 40x40 pixels by use of an image scaling technique. For each of the nodules, the same scaling factor was applied for scaling of the corresponding test candidates, which had the same center location with the corresponding template, and whose matrix size was 48x48 pixels.

[0064] Each template was then right/left reversed, rotated by two angles of -10 and +10 degrees, and scaled by two factors of 0.9 and 1.1 in order to correct the inaccuracy of the estimated nodule size by the radiologists. By combining the three techniques for increasing of the number of templates, we obtained 180 (18x10) templates for malignant nodules and

WO 02/43562

PCT/US01/43146

180 (18x10) templates for benign nodules. The background trend was corrected for all the templates and the test candidates prior to the determination of cross correlation values between them. A round robin test method was employed so that a test candidate and the corresponding derived templates were not included in the template set used for testing. [0065] Figure 18 shows the maximum cross-correlation values for 10 malignant nodules (circles) and 10 benign nodules (dots) obtained with the multiple-templates matching technique by use of 180 (18x10) malignant nodule templates and 180 (18x10) benign nodule templates. It is apparent in Fig. 18 that malignant nodules tend to be located above the diagonal line, whereas benign nodules tend to be located below the diagonal line, which indicates that malignant nodules are generally more similar to malignant nodule templates, and benign nodules are generally more similar to benign nodule templates. If the diagonal line is used as a threshold, that is, if those candidates above the diagonal line are accepted as malignant nodules and those candidates below the diagonal line as benign nodules, then it is possible to eliminate 8 (80%) benign nodules while retaining all the malignant nodules. This result indicates that the multiple-templates matching technique has the ability to distinguish benign nodules from malignant nodules, in order to improve the performance of CAD schemes for classification between benign lesions and malignant lesions due to many cancers, such as lung cancer, breast cancer, colon cancer, and stomach cancer.

#### Computer Program Product

[0066] The mechanisms and processes set forth in the present description may be implemented using a conventional general purpose microprocessor or computer programmed according to the teachings in the present specification, as will be appreciated by those skilled in the relevant art(s). Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will also be apparent to those skilled in the relevant art(s). However, as will be readily apparent to those skilled in the art, the present invention also may be implemented by the preparation of application-specific integrated circuits or by interconnecting an appropriate network of conventional component circuits.

WO 02/43562

PCT/US01/43146

[0067] The present invention thus also includes a computer-based product which may be hosted on a storage medium and include instructions which can be used to program a general purpose microprocessor or computer to perform processes in accordance with the present invention. This storage medium can include, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, magneto-optical disks, ROMs, RAMs, EPROMs, EEPROMs, flash memory, magnetic or optical cards, or any type of media suitable for storing electronic instructions.

[0068] The programming of the general purpose microprocessor or computer may include a software module for digitizing and storing images obtained from an image acquisition device (not shown). Alternatively, the present invention can also be implemented to process digital data derived from images obtained by other means, such as a picture archive communication system (PACS) or directly from an imaging device which produces digital image data. In other words, the digital images being processed may already be in existence in digital form and need not be converted to digital form in practicing the invention.

[0069] Numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore to be understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described herein.

WO 02/43562

PCT/US01/43146

## List of References:

- [0070] 1. J. V. Forrest and P. J. Friedman, "Radiologic errors in patients with lung cancers," *West J. Med.* 134, 485-490, (1981).
- [0071] 2. D. P. Naidich, F. A. Zerhouni, and S. S. Siegelman, *Computer tomography of thorax* (Raven, New York, 1984), pp. 171-199.
- [0072] 3. J. R. Muhm, R. S. Miller, R. S. Fontana, et al, "Lung cancer detected during a screening program using four-month chest radiographs," *Radiology* 148, 609-615 (1983).
- [0073] 4. M. L. Giger, K. Doi, and H. MacMahon, "Image feature analysis and computer-aided diagnosis in digital radiography. 3. Automated detection of nodules in peripheral lung fields," *Med. Phys.* 15, 158-166 (1988).
- [0074] 5. M. L. Giger, K. Doi, H. MacMahon, C. E. Metz, and F. F. Yin, "Pulmonary nodules: computer aided detection in digital chest image," *RadioGraphics* 10, 41-51 (1990).
- [0075] 6. T. Matsumoto, H. Yoshimura, K. Doi, M. L. Giger, A. Kano, H. MacMahon, K. Abe, and S. M. Montner, "Image feature analysis of false-positive diagnoses produced by automated detection of lung nodules," *Invest. Radiol.* 27, 587-597 (1992).
- [0076] 7. Y. C. Wu, K. Doi, M. L. Giger, C. E. Metz, and W. Zhang, "Reduction of false positives in computerized detection of lung nodules in chest radiographs using artificial neural networks, discriminant analysis and a rule-based scheme," *J. Digital Imag.* 7, 196-207 (1994).
- [0077] 8. T. Kobayashi, X. W. Xu, H. MacMahon, C. E. Metz, and K. Doi, "Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs," *Radiology* 199, 843-848 (1996).
- [0078] 9. X. W. Xu, K. Doi, T. Kobayashi, H. MacMahon, and M. L. Giger, "Development of an improved CAD scheme for automated detection of lung nodules in digital chest images," *Med. Phys.* 24, 1395-1403 (1997).
- [0079] 10. J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database

WO 02/43562

PCT/US01/43146

for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules," *AJR* 147, 71-74 (2000).

- [0080] 11. X. W. Xu, K. Doi, "Image feature analysis for computer-aided diagnosis: Accurate determination of ribcage boundary in chest radiographs," *Med. Phys.* 22, 617-626 (1995).
- [0081] 12. X. W. Xu, K. Doi, "Image feature analysis for computer-aided diagnosis: Detection of right and left hemidiaphragm edges and delineation of lung field in chest radiographs," *Med. Phys.* 23, 1613-1624 (1996).
- [0082] 13. S. G. Armato III, M. L. Giger, C. J. Moran, J. T. Blackburn, K. Doi, and H. MacMahon, "Computerized detection of pulmonary nodules on CT scans," *Radiographics* 19, 1303-1311, (1999).

WO 02/43562

PCT/US01/43146

**CLAIMS:**

[0083] 1. In a method to determine whether a candidate abnormality in a medical digital image is an actual abnormality, the improvement comprising:

obtaining plural first templates and plural second templates respectively corresponding to predetermined abnormalities and predetermined non-abnormalities;  
comparing the candidate abnormality in the medical digital image with the obtained first and second templates to derive cross-correlation values between the candidate abnormality and each of the obtained first and second templates;

determining the largest cross-correlation value derived in the comparing step and whether the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates or with the second templates; and

determining the candidate abnormality to be an actual abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the first templates and determining the candidate abnormality to be a non-abnormality when the largest cross-correlation value is produced by comparing the candidate abnormality with the second templates.

[0084] 2. The method of Claim 1, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining plural first templates and plural second templates which are smaller in size than the candidate abnormality.

[0085] 3. The method of Claim 2, wherein said comparing step comprises:

comparing the candidate abnormality with the obtained first and second templates to derive cross-correlation values between the candidate abnormality and each of the obtained first and second templates; and

shifting the first and second templates in relation to said candidate abnormality to derive cross-correlation values between the candidate abnormality and each of the shifted first and second templates.

[0086] 4. The method of Claim 1, wherein said step of obtaining plural first templates and plural second templates comprises:

WO 02/43562

PCT/US01/43146

producing additional first and second templates which are mirror images of the obtained plural first templates and plural second templates.

[0087] 5. The method of Claim 1, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0088] 6. The method of Claim 1, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0089] 7. The method of Claim 2, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are mirror images of the obtained plural first templates and plural second templates.

[0090] 8. The method of Claim 2, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0091] 9. The method of Claim 2, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0092] 10. The method of Claim 3, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are mirror images of the obtained plural first templates and plural second templates.

[0093] 11. The method of Claim 3, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.



WO 02/43562

PCT/US01/43146

[0094] 12. The method of Claim 3, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0095] 13. The method of Claim 7, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0096] 14. The method of Claim 7, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0097] 15. The method of Claim 10, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0098] 16. The method of Claim 10, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0099] 17. The method of Claim 1, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;

comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure; and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

WO 02/43562

PCT/US01/43146

[0100] 18. The method of Claim 2, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;

comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure; and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0101] 19. The method of Claim 3, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;

comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure; and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0102] 20. The method of Claim 16, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;

comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure; and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

WO 02/43562

PCT/US01/43146

[0103] 21. The method of claim 1, further comprising:

obtaining plural third templates and plural fourth templates respectively corresponding to predetermined malignant abnormalities and predetermined benign abnormalities;

comparing the actual abnormality with the obtained third and fourth templates to derive cross-correlation values between the actual abnormality and each of the obtained third and fourth templates;

determining the largest cross-correlation value derived in the preceding comparing step and whether the largest cross-correlation value is produced by comparing the actual abnormality with the third templates or with the fourth templates; and

classifying the actual abnormality as a malignant abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the third templates and classifying the actual abnormality as a benign abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the fourth templates.

[0104] 22. The method of claim 2, further comprising:

obtaining plural third templates and plural fourth templates respectively corresponding to predetermined malignant abnormalities and predetermined benign abnormalities;

comparing the actual abnormality with the obtained third and fourth templates to derive cross-correlation values between the actual abnormality and each of the obtained third and fourth templates;

determining the largest cross-correlation value derived in the preceding comparing step and whether the largest cross-correlation value is produced by comparing the actual abnormality with the third templates or with the fourth templates; and

classifying the actual abnormality as a malignant abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the third templates and classifying the actual abnormality as a benign abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the fourth templates.

[0105] 23. The method of claim 3, further comprising:

WO 02/43562

PCT/US01/43146

obtaining plural third templates and plural fourth templates respectively corresponding to predetermined malignant abnormalities and predetermined benign abnormalities;

comparing the actual abnormality with the obtained third and fourth templates to derive cross-correlation values between the actual abnormality and each of the obtained third and fourth templates;

determining the largest cross-correlation value derived in the preceding comparing step and whether the largest cross-correlation value is produced by comparing the actual abnormality with the third templates or with the fourth templates; and

classifying the actual abnormality as a malignant abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the third templates and classifying the actual abnormality as a benign abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the fourth templates.

[0106] 24. The method of claim 17, further comprising:

obtaining plural third templates and plural fourth templates respectively corresponding to predetermined malignant abnormalities and predetermined benign abnormalities;

comparing the actual abnormality with the obtained third and fourth templates to derive cross-correlation values between the actual abnormality and each of the obtained third and fourth templates;

determining the largest cross-correlation value derived in the preceding comparing step and whether the largest cross-correlation value is produced by comparing the actual abnormality with the third templates or with the fourth templates; and

classifying the actual abnormality as a malignant abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the third templates and classifying the actual abnormality as a benign abnormality when the largest cross-correlation value is produced by comparing the actual abnormality with the fourth templates.

[0107] 25. In a method of classifying an abnormality in a medical digital image, the improvement comprising:

WO 02/43562

PCT/US01/43146

obtaining plural first templates and plural second templates respectively corresponding to predetermined malignant abnormalities and predetermined benign abnormalities;

comparing the abnormality in the medical digital image with the obtained first and second templates to derive cross-correlation values between the abnormality and each of the obtained first and second templates;

determining the largest cross-correlation value derived in the comparing step and whether the largest cross-correlation value is produced by comparing the abnormality with the first templates or with the second templates; and

classifying the abnormality as a malignant abnormality when the largest cross-correlation value is produced by comparing the abnormality with the first templates and classifying the abnormality as a benign abnormality when the largest cross-correlation value is produced by comparing the abnormality with the second templates.

[0108] 26. The method of Claim 25, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining plural first templates and plural second templates which are smaller in size than said abnormality in said medical digital image.

[0109] 27. The method of Claim 26, wherein said comparing step comprises:

comparing the candidate abnormality with the obtained first and second templates to derive cross-correlation values between the candidate abnormality and each of the obtained first and second templates; and

shifting the first and second templates in relation to the abnormality to derive cross-correlation values between the abnormality and each of the shifted first and second templates.

[0110] 28. The method of Claim 25, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are mirror images of the obtained plural first templates and plural second templates.

[0111] 29. The method of Claim 25, wherein said step of obtaining plural first templates and plural second templates comprises:

WO 02/43562

PCT/US01/43146

producing additional first templates which are scaled versions of the obtained plural first templates.

[0112] 30. The method of Claim 25, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0113] 31. The method of Claim 26, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are mirror images of the obtained plural first templates and plural second templates.

[0114] 32. The method of Claim 26, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0115] 33. The method of Claim 26, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0116] 34. The method of Claim 27, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are mirror images of the obtained plural first templates and plural second templates.

[0117] 35. The method of Claim 27, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0118] 36. The method of Claim 27, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

WO 02/43562

PCT/US01/43146

[0119] 37. The method of Claim 34, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first templates which are scaled versions of the obtained plural first templates.

[0120] 38. The method of Claim 35, wherein said step of obtaining plural first templates and plural second templates comprises:

producing additional first and second templates which are rotated versions of the obtained plural first templates and plural second templates.

[0121] 39. The method of Claim 25, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;

comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure; and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0122] 40. The method of Claim 26, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;

comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure; and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0123] 41. The method of Claim 27, wherein said step of obtaining plural first templates and plural second templates comprises:

WO 02/43562

PCT/US01/43146

obtaining candidate first templates and candidate second templates;  
comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure;  
and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0124] 42. The method of Claim 34, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;  
comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure;  
and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0125] 43. The method of Claim 37, wherein said step of obtaining plural first templates and plural second templates comprises:

obtaining candidate first templates and candidate second templates;  
comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure;  
and

comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0126] 44. The method of Claim 38, wherein said step of obtaining plural first templates and plural second templates comprises:



WO 02/43362

PCT/US01/43146

obtaining candidate first templates and candidate second templates;  
comparing the candidate first templates to at least one known non-abnormality structure and retaining as first templates only those candidate first templates which exhibit a predetermined degree of non-similarity to the at least one known non-abnormality structure;  
and  
comparing the candidate second templates to at least one known abnormality structure and retaining as second templates only those candidate second templates which exhibit a predetermined degree of non-similarity to the at least one known abnormality structure.

[0127] 45. A system for implementing the method of any one of claims 1-44.

[0128] 46. A computer readable medium storing a program for performing the steps recited in any one of claims 1-44.

WO 02/43562

1/22

PCT/US01/43146

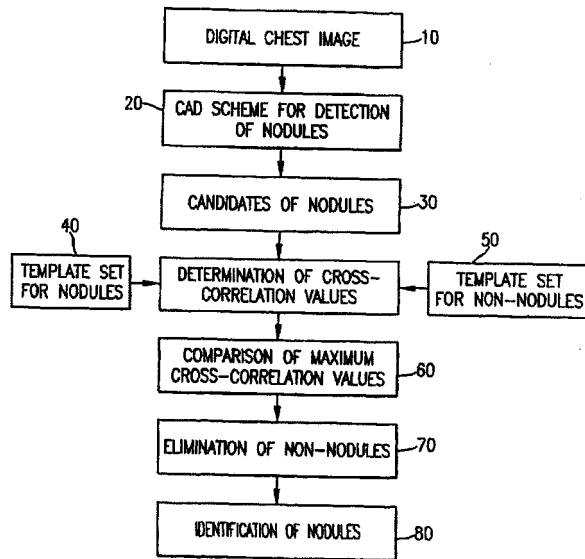


FIG.1

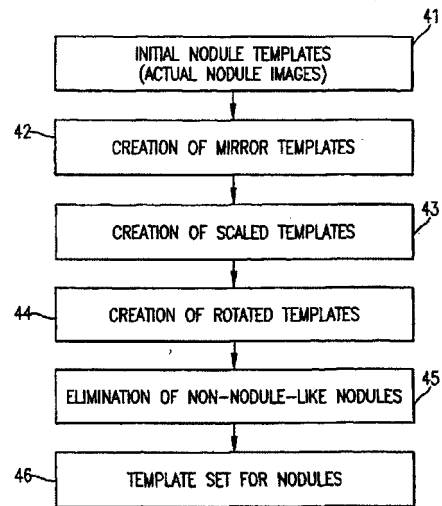


FIG.2

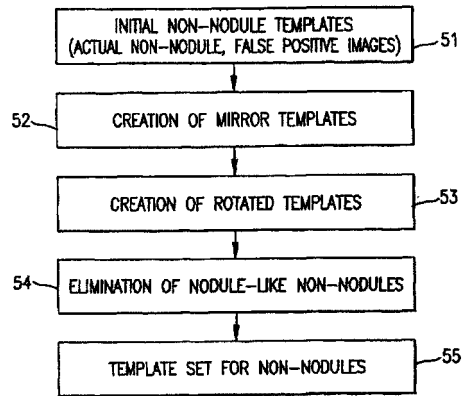


FIG.3

WO 02/43562

PCT/US01/43146

4/22



FIG. 4



FIG. 5

WO 02/43562

PCT/US01/43146

5/22



FIG. 6



FIG. 7

WO 02/43562

6/22

PCT/US01/43146

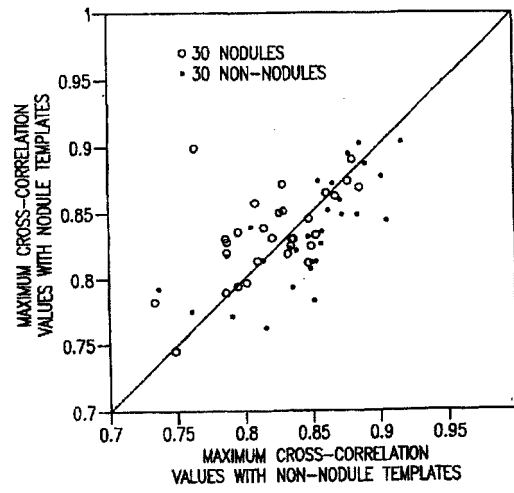


FIG.8

WO 02/43562

7/22

PCT/US01/43146

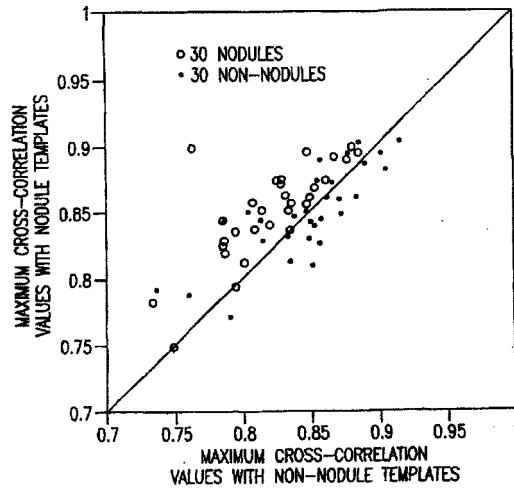


FIG.9A



WO 02/43562

8/22

PCT/US01/43146

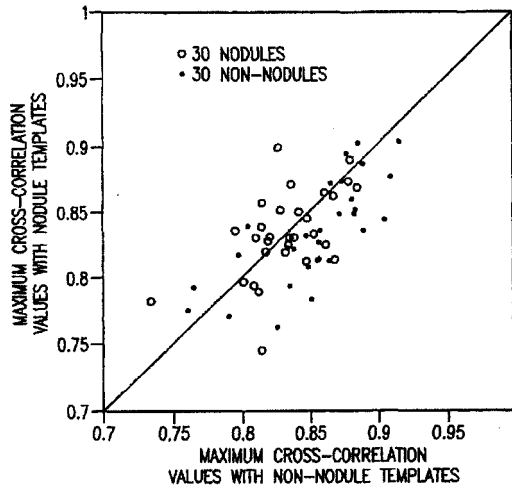


FIG.9B

WO 02/43562

9/22

PCT/US01/43146

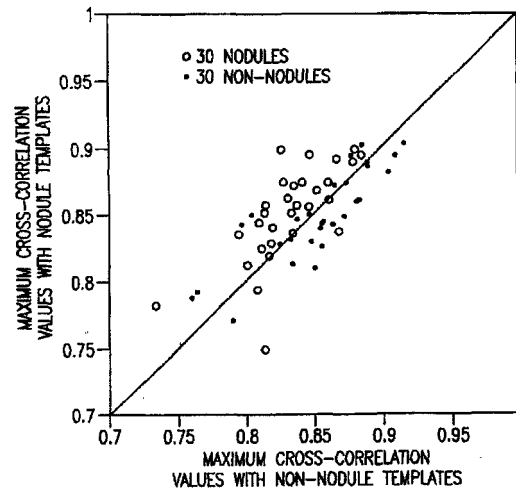


FIG.9C

WO 02/43562

10/22

PCT/US01/43146

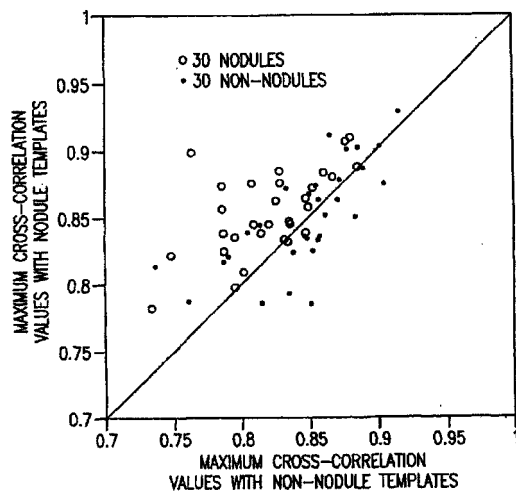


FIG.10

WO 02/43562

11/22

PCT/US01/43146

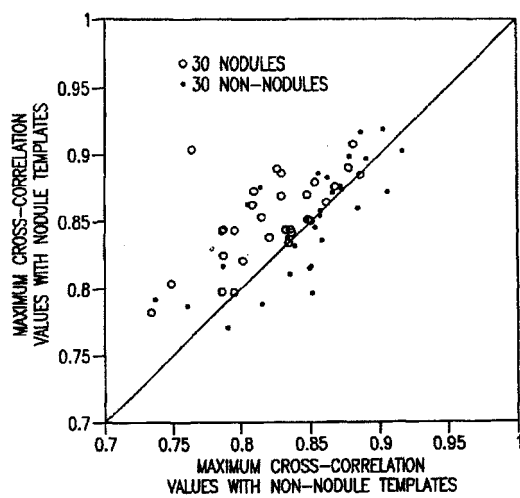


FIG.11A

WO 02/43562

12/22

PCT/US01/43146

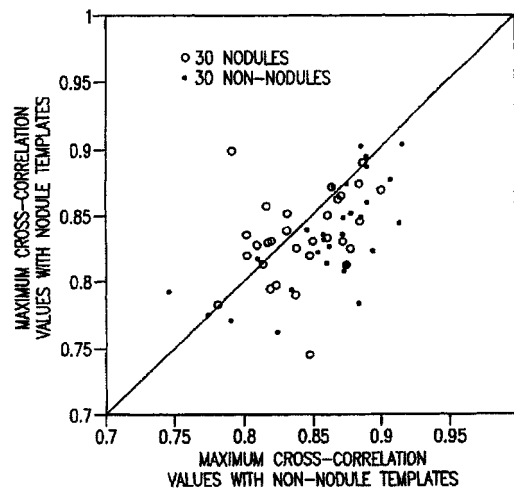


FIG.11B

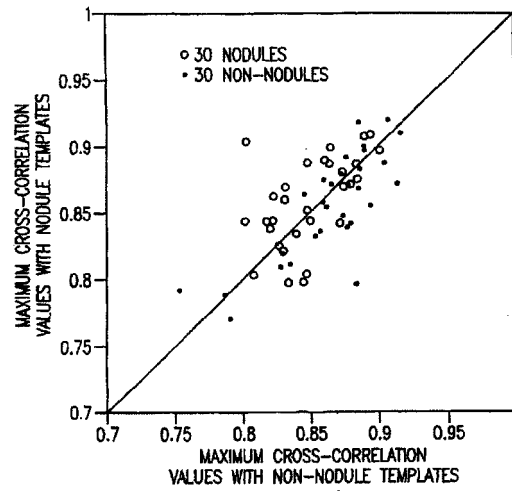


FIG.11C

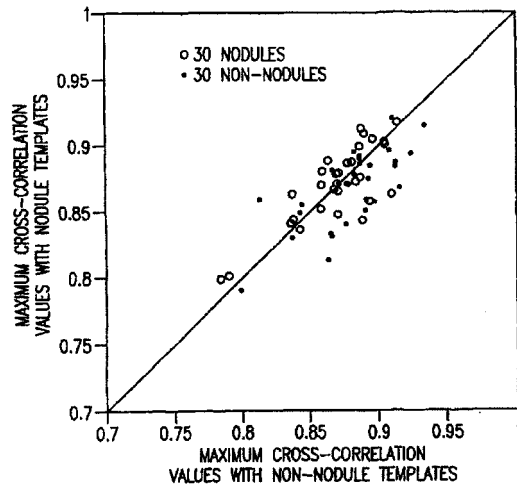


FIG.12A

WO 02/43562

15/22

PCT/US01/43146

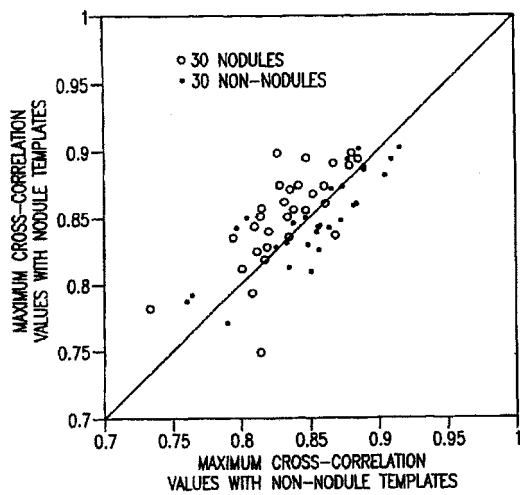


FIG.12B



WO 02/43562

16/22

PCT/US01/43146

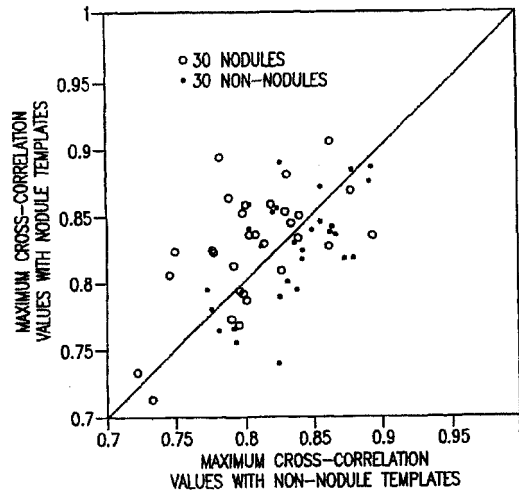


FIG.12C

WO 02/43562

PCT/US01/43146

17/22



FIG.13

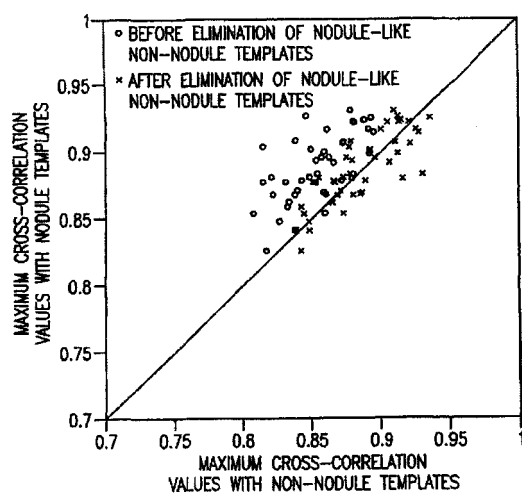


FIG.14

WO 02/43562

19/22

PCT/US01/43146

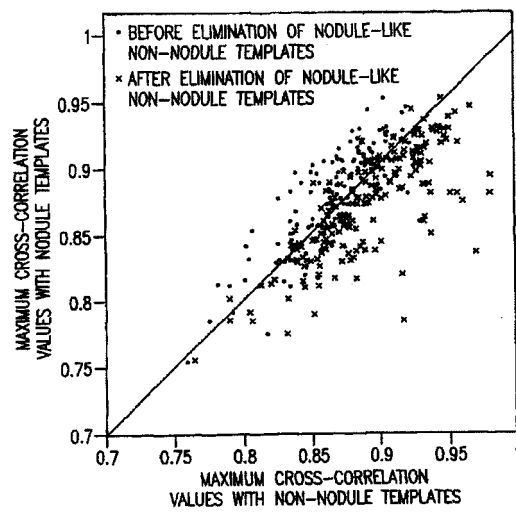


FIG.15

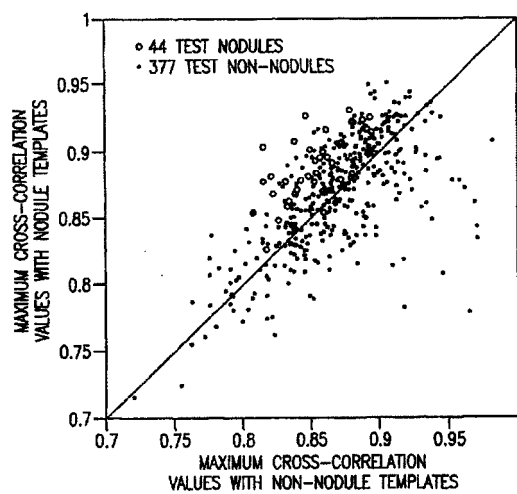


FIG.16

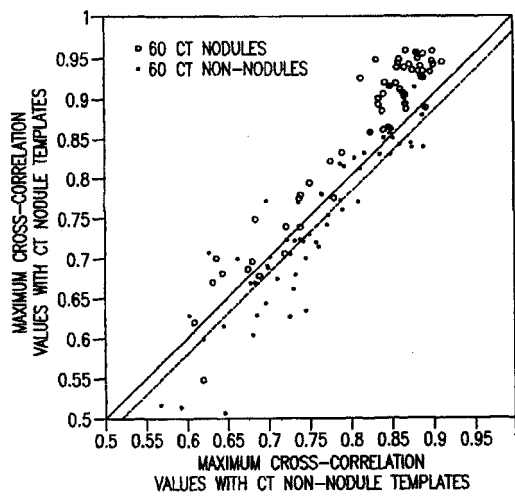


FIG.17

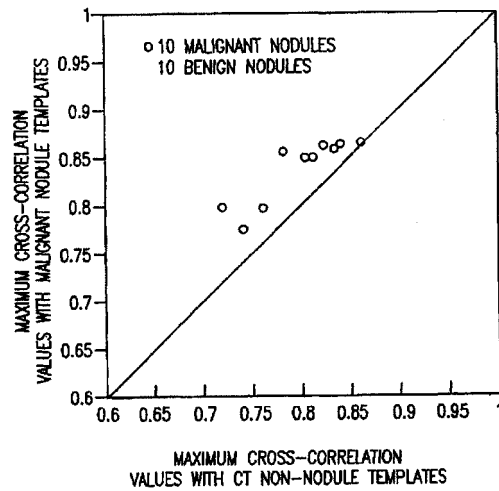


FIG.18

## フロントページの続き

(51)Int.Cl. <sup>7</sup>	F I	テーマコード (参考)
G 0 1 R 33/32	A 6 1 B 6/00	3 5 0 D
G 0 6 T 1/00	A 6 1 B 5/05	3 8 0
G 0 6 T 7/00	G 0 1 N 24/02	5 2 0 Y

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TC),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,OM,PH,P L,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZW

(特許庁注：以下のものは登録商標)

フロッピー

(74)代理人 100109830

弁理士 福原 淑弘

(74)代理人 100084618

弁理士 村松 貞男

(74)代理人 100092196

弁理士 橋本 良郎

(72)発明者 リー、キアン

アメリカ合衆国、イリノイ州 6 0 5 1 4 クラレンドン・ヒルズ、ナンバー 1 5、コンコルド・  
レーン 5 7 3 0

(72)発明者 桂川 茂彦

大分県大分市望みが丘 2 1 - 8

(72)発明者 土井 邦雄

アメリカ合衆国、イリノイ州 6 0 5 2 1 ウィローブルック、レーン・コート 6 4 1 5

F ターム(参考) 4C093 AA26 CA50 DA01 DA03 DA06 FD03 FD20 FF12 FF17 FF19

FF20 FF31 FF34

4C096 AB38 AC04 AD14 DC20 DC21 DC22

4C601 EE30 JC05 JC06 JC07 JC08

5B057 AA07 BA03 BA05 BA07 DA03 DB02 DB05 DB09 DC34

5L096 BA06 BA13 EA43 FA34 FA67 GA34 HA07 JA03 JA09

# 【要約の続き】

相関値が候補異常を第2テンプレートと比較することによって生成された場合には候補異常を非異常と決定することを含む。真の異常は同様に、悪性および良性異常に対応する追加テンプレートと比較することによって得られるさらなる相互相関値に基づいて、悪性または良性と分類される。